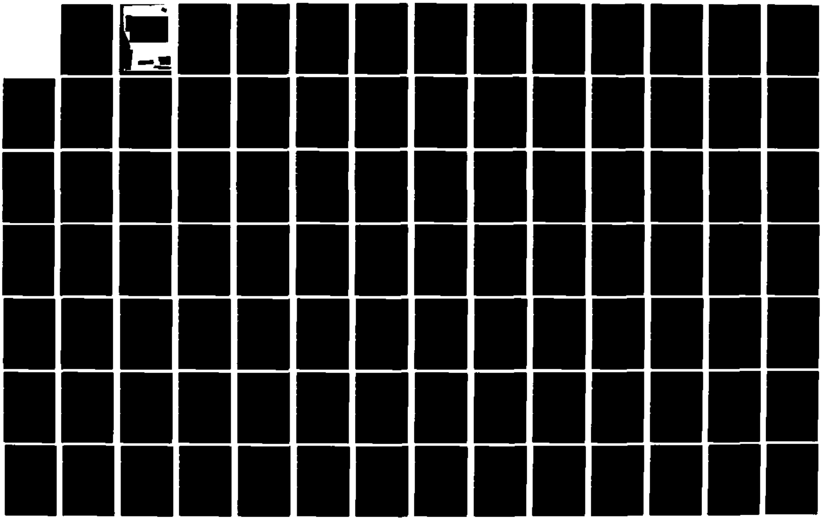
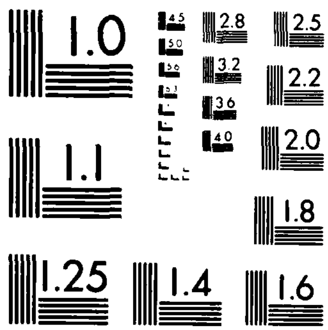
AD-A149 514 BLUNDER DETECTION USING A SEQUENTIAL LEAST SQUARES ADJUSTMENT(U) VIRGINIA POLYTECHNIC INST AND STATE UNIV BLACKSBURG D A CAMPBELL NOV 84 1/2

UNCLASSIFIED F/G 12/1 NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963 A

BlUNDER DETECTION USING A SEQUENTIAL LEAST SQUARES

ADJUSTMENT

by

Douglas A. Campbell

Thesis submitted to the Graduate Facility of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

in

Civil Engineering

BlUNDER DETECTION USING A SEQUENTIAL LEAST SQUARES

ADJUSTMENT

by

Douglas A. Campbell

Thesis submitted to the Graduate Facility of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

in

Civil Engineering

APPROVED:

Steven D. Johnson, Chairman

Charles Schwarz

James Baker

November, 1984

Blacksburg, Virginia

BLUNDER DETECTION USING A SEQUENTIAL LEAST SQUARES ADJUSTMENT

by

Douglas A. Campbell

Blunder detection is a key step in achieving optimal results from a least squares adjustment. One approach to blunder detection is to employ a detection algorithm in a sequential least squares adjustment. A sequential adjustment technique compares the variability of the observations added sequentially to that of the observations already incorporated into the solution. The rejection criterion for observations is based upon an F-statistic. The F-statistic is formed by the ratio of key, independent quadratic forms of residuals corresponding to the observations added sequentially and those already accepted in the solution. Neither an a priori nor an a posteriori reference variance is required for the blunder detection algorithm. The sequential algorithm is computationally efficient because it combines the least squares adjustment and the blunder detection into one step. The sequential least squares blunder detection algorithm is programmable and is tested on measured data sampled from a simple model, a straight line.

ACKNOWLEDGEMENTS

I wish to thank Dr. Charles Schwarz for his help, advise, and encouragement during the development of this thesis.

Special appreciation goes also to Dr. Steven D. Johnson for his guidance throughout the academic year.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. Introduction

1.1 Problem Definition

Blunder detection is a key step in achieving optimal results from a least squares adjustment. For the purpose of this paper, a blunder is any gross error in the observation data or adjustment procedure that causes the adjustment to be invalid for its intended purpose. Using this definition, blunder detection can be considered as a measure of quality control or reliability of the adjustment results. Modern developments in blunder detection began with Baarda , who developed a method called data snooping for detecting gross errors in geodetic nets.(Baarda, 1967) Baarda treats gross error detection as a statistical reliability problem and he offers several objective measures for quality of results obtained from least squares adjustments. Many have refined his techniques and have adapted them to aerotriangulation. Pope also developed a theoretical basis for detecting bad data using the "Tau" criterion on the results of simultaneous least squares adjustments of triangulation networks. (Pope, 1976) These methods provide a more sound basis for data rejection than was previously available and have provided many new topics for study. One such topic is the application

of statistically rigorous gross error detection methods to sequential least squares adjustments.

There are many advantages in developing a blunder detection method incorporated in the sequential least squares adjustment technique.

1. Sequential techniques permit deletion of bad data with greater computational efficiency than batch adjustments.
2. Many problems are better suited for a sequential adjustment.
3. A successful sequential blunder check provides the blunder's location.
4. Use of a sequential technique implies that after each iteration there is at most one blunder , the most recently added observation.
5. One may examine the effect of adding and deleting the observations on parameter estimates and their covariance matrix.
6. Reobservation may be impossible after simultaneous adjustment. The sequential approach may allow re-observation and thus eliminate singularities in the normal equation coefficient matrix caused by deleting blunders after a simultaneous adjustment.
7. A remeasurement capability may allow achievement of optimum levels for type I and II errors.

There are also several challenges to approaching blunder detection with a sequential algorithm. First, sequential adjustments require the inverse of the normal equation coefficient matrix. Inversion is not an additional burden since most statistically based blunder detection methods also require the inverse. The sequential technique is actually computationally more efficient because it combines the blunder detection step with the actual adjustment. One must also be aware that many techniques already developed rely on the central limit theorem to employ a number parametric tests. Sequential analysis would require analysis of the effects of sample size and redundancy on the test. Then, there is the fact that the sequential technique does not yield the true residual at a given step in the adjustment, but a residual based on parameters estimated from observations added to that point. A new estimate of the reference variance will be required for each step of the sequential adjustment. Finally, the assumption that only one blunder is present may be violated as the test sensitivity increases when the observations are added. In other words, as the solution begins to converge when more observations are added, is it possible that previously undetectable blunders could, in fact, be detected? The answer to these questions will be a major portion of this thesis.

1.2 Thesis Objective

The objective of this thesis is to develop and test an algorithm for detecting blunders in observation data using a sequential least squares adjustment technique. The objective will also entail developing a computer program as the means of testing the algorithm. The algorithm should meet the following requirements:

1. It should detect blunders of small magnitude and blunders of large magnitude equally well.
2. It should provide blunder location , if possible, without reobservation.
3. The algorithm should be computationally efficient, and should not be restricted to standard cases.
4. It must determine tolerances and the maximum undetectable blunder.
5. The algorithm should not assume a priori knowledge of the blunder's magnitude, location, or frequency of occurence.
6. It should incorporate statistical measures of confidence for types I and II errors.

1.3 Definition of a Blunder

Before proceding with the algorithm development, it would be useful to define blunder. A blunder is an outlier, or observation which appears to be inconsistent with the remainder of the data set.(Barnett and Lewis, 1978). Outliers, though, are not necessarily bad or erroneous and should not universally be rejected when found. Their presence may help explain, or at least help identify, a true anomaly that should actually be incorporated into the solution. This requires a thorough knowledge of the model or system under study before embarking on any technique that will automatically reject data behaving as outliers. There is inherent variability in any measuring process. However, this variablity, which is the combination of many small random errors can be modeled statistically. Any data, then, that do not conform to the statistical model are blunders. These blunders may arise from many souces such as recording errors, data transmission errors, or point misidentification, and their magnitude may range from very small to very large. It is useful to divide blunders into two categories. First, there are large blunders that are detectable without having to resort to rigorous statistical methods as suggested by El-Hakim. (El-Hakim and Zeiman, 1984) Second, there are blunders of

2. Algorithm Proposal

2.1 Assumptions

There are several assumptions required for developing a blunder detection scheme based upon the examination of residuals from a least squares adjustment. First, there are four assumptions concerning the observational errors.

1. $E(e_i) = 0$ which implies that $E(v_i) = 0$ as well. This is demonstrated in Appendix J.
2. $Var(e_i) = \sigma_o^2$ where e_i is the error of the observation of unit weight. σ_o^2 will hereafter be refered to as the reference variance. This implies that $Var(V) = \sigma_o^2 \bar{Q}$, where V is the vector of residuals.
3. $Cov(e_i, e_j) = 0$ for $i \neq j$, which means that there is no correlation among the observational errors.
4. $e_i \sim N(0, \sigma_o^2)$ which means that the individual residuals are also distributed as a normal, V is distributed as a multivariate normal, and parametric statistical tests may be used.

These four assumptions provide a basis for using residuals for blunder detection due to the relationship between the

small magnitude which are not readily obvious but can cause the solution to be invalid for its intended purpose. Blunders in this category are very difficult to indentify without statistical analysis and will be the primary target of the algorithm to be developed. It follows that any technique for detecting blunders of the second category would certainly work on large blunders overlooked by a non-statistical method. A blunder, then, is an observation that does not behave statistically as one would expect regardless of the cause. The question of what to do with blunders, once found, will be discussed later in this paper.

1.4 Thesis Organization

The remainder of this paper is organized to accomplish the thesis objective. Chapter 2 will propose an algorithm for blunder detection using a sequential least squares adjustment procedure. Chapter 3 will present an evaluation of the algorithm and Chapter 4 will provide the results of how the algorithm performed on a simple data set. Finally, Chapter 5 will summarize how well the algorithm satisfies the requirements in Section 1.2 and offers questions requiring further study. The theoretical basis for the algorithm is developed in the Appendices and throughout the text; equations will be introduced that are fully developed in the appendices. Also, Appendix A contains a list of the symbols used in the least squares model presented in this thesis.

residuals and observational errors developed in Appendix J. There are additional assumptions required which concern the observations. These assumptions will be necessary when a sequential adjustment method will be used.

1. The weight matrix of the observations is block diagonal. This implies that individual observations are independent, or at least the observations may be divided into small independent blocks. Observations can then be accepted or rejected independently without any inference concerning other observations.
2. The total number of blunders in the observation data is small and the redundancy exceeds , by far , the number of blunders. If this is not the case, the set of observations reflects problems of a different nature such as data transmission or modeling errors.
3. The observations, less any blunders, should form a well conditioned algebraic system. Furthermore, at least (u + r1) observations must agree with assumed stochastic properties and form the well conditioned algebraic system. The variables u and r1 are the number of parameters and initial redundancy. There may be no systemmatic errors and there must be a correct initial solution from which the sequential adjustment may proceed.
4. The observations to be added 'p' at a time during a

sequential adjustment can be added in a manner such that the rank of the cofactor matrix for residuals corresponding to the 'p' observations is p. This is important because the algorithm to be developed will require that $\bar{Q}_{22}$ and $[Q_2 + B_2 N_1^{-1} B_2^t]$ be inverted.

All of these assumptions, including the assumption of normality of the observational errors, are reasonable for most measuring systems. Anyone applying a blunder detection algorithm during an adjustment should have studied the system sufficiently to verify the validity of these assumptions.

2.2 Key Relationships in Sequential Adjustments

A least squares solution from observation equations being added sequentially will produce results identical to those of other approaches. However, the observation equation method lends itself to very simple partitioning and permits working with the observations directly and individually. Appendices B and C show how a set of observation equations can be partitioned into a base group and a test

group. The resulting residuals are then

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} F_1 \\ F_2 \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} X_2 . \qquad 2.1$$

These residuals can be estimated using only the observations in L_1.

$$\begin{bmatrix} \hat{V}_1 \\ \hat{V}_2 \end{bmatrix} = \begin{bmatrix} F_1 \\ F_2 \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} X_1 \qquad 2.2$$

$\hat{V}_1$ and $\hat{V}_2$ are functions of only the observations in the base group while the V_1 and V_2 are functions of the observations in both the base and test groups. $\hat{V}_1$ and $\hat{V}_2$, though, may be expressed as functions of V_1 and V_2 by

$$\hat{V}_1 = V_1 - \bar{Q}_{12} \bar{Q}_{22}^{-1} V_2 \qquad 2.3$$

$$\hat{V}_2 = Q_2 \bar{Q}_{22}^{-1} V_2 \qquad 2.4$$

These two equations are very useful because they permit the total sum of weighted, squared residuals to be computed for each step of the sequential adjustment as the sum of the effects of the base group and the effects of the test group. From Appendix G,

$$V^t W V = \hat{V}_1^t W_1 \hat{V}_1 + V_2^t \bar{Q}_{22}^{-1} V_2 \qquad 2.5$$

This is a key relationship because the residual estimates corresponding to the observations in the base group are actually the residuals for the same observations prior to adding the 'p' observations of the test group sequentially. The total sum of weighted, squared residuals can be computed for each step of the sequential adjustment and carried forward for the next group of 'p' observations. Appendix H explains also that the two quadratic forms on the right side of equation 2.5 are independent and distributed as CHI-SQUARE, which follows the assumption that the observational errors and, hence, the residuals are normaly distributed. These two distributions are

$$\hat{V}_1^t W_1 \hat{V}_1 \sim \text{CHI-SQUARE}(r1) \times \sigma_0^2 \qquad 2.6$$

$$V_2^t \bar{Q}_{22}^{-1} V_2 \sim \text{CHI-SQUARE}(p) \times \sigma_0^2 \qquad 2.7$$

where r1 and p are the degrees of freedom of the two CHI-SQUARE distributions.

The two quadratic forms will be the basis for a blunder detection algorithm. However, it is first necessary to determine whether an a priori or a posteriori reference variance should be used.

2.3 A Priori Versus A Posteriori Reference Variance

There are two distinct approaches to blunder detection that depend on whether or not an a priori reference variance is available. Stefanovic is an advocate of using an a priori variance (Stefanovic, 1981). He assumes that the mathematical models and measurements systems are sufficiently tested in advance to eliminate systemmatic errors and determine an a priori reference variance. He also points out that the a posteriori variance can always be tested against the a priori variance using a simple F-test. He finds that the use of an a posteriori variance in applications such as Pope's Tau distribution (Pope, 1976) has a serious disadvantage. A blunder among the observations inflates the estimated variance and makes any statistical test incorporating an a posteriori estimate less likely to reject blunders of a magnitude small enough to have been overlooked by nonstatistical methods. His dual CHI-SQUARE test (Stefanovic, 1981) examines both the base group and the test group using equations 2.6 and 2.7.

Pope (Pope 1976), however, argues that in reality, the actual variance and the a priori variance differ significantly. Often, the a priori variance is too small and causes too many observations to be rejected. The ideal algorithm for blunder detection, then, should not require either reference variance, and thus eliminate the apparent weaknesses of both of these approaches.

2.4 The Proposed Test

One possible test for blunders that does not require incorporating either an a priori or a posteriori reference variance is an F-test. Because the two quadratic forms (equations 2.6 and 2.7) are independently distributed as CHI-SQUARE, by definition the test statistic

$$T.S. = \frac{[(V_2^t \bar{Q}_{22}^{-1} V_2)/p]}{[(\hat{V}_1^t W_1 \hat{V}_1)/r1]} \qquad 2.8$$

is distributed as an F with p and r1 degrees of freedom. The test statistic is compared to a critical value defined by the F statistic with p and r1 degrees of freedom at a given level

of significance. The 'p' observations of the test group are rejected if the test statistic exceeds the F statistic at the required level of significance. The test simply determines whether the observations in the test group have more variability than those of the base group. Rejection occurs when the test group does demonstrate statistically greater variability than the base group. This test has validity when the observational errors and corresponding residuals have expected values of zero. The test is also independent of any a priori variance, and simply requires that consistent accuracy and precision in the observations be maintained. The statistical test should proceed as follows:

Ho : The observations of the test group have a variance less than or equal to that of the base group. Or,

$$\sigma^2_{test} \leq \sigma^2_{base}$$

Ha : The observations of the test group have a variance greater than that of the base group. Or,

$$\sigma^2_{test} > \sigma^2_{base}$$

Reject Ho if T.S. > $F(\alpha, p, rl)$.

Even though this test does not rely on an a priori variance, it does require that the initial solution be statistically and algebraically acceptable. The requirement for a good initial solution will be discussed later but should be

mentioned now because the sequential solution must build upon an initial solution. If there is a requirement to satisfy an a priori reference variance, this test still permits the a posteriori estimate to be tested using an F-test. The advantage, though, is that the proposed sequential technique does not allow the statistically detectable blunders to inflate the estimate of the variance. The proposed test provides a truer picture of the variability in the measurement system and a better comparison of the a priori and a posteriori reference variances when such a comparison is required.

3. A Critique of the Algorithm

3.1 Advantages of A Sequential Algorithm

The proposed algorithm has a number of strengths or advantages when compared to many simultaneous adjustment blunder detection methods. First, it has a sound theoretical basis. Unlike many techniques, the F-test of this algorithm does not rely on an arbitrary 3-sigma cutoff or similar artificial critical value. It also does not rely on a restrictive assumption that the data can contain at most one blunder as do methods such as rejecting the observation with the largest residual or Baarda's data snooping. Another theoretical advantage is that this test does not rely on either an a priori or a posteriori variance. The sequential method rejects blunders before they can contaminate the estimate of the reference variance. The standard F-test is, then, a better indication of how the actual measurement system is behaving relative to an a priori variance if such a comparison is required.

Aside from these theoretical advantages over traditional approaches to blunder detection, this algorithm also has computational and practical advantages over those simul-

taneous adjustment methods that are theoretically sound. The sequential adjustment offers much greater computational efficiency in several ways. In cases where more than one blunder may be present, blunder location following a batch adjustment requires a brute search, while the sequential method provides location at the same time that a blunder is detected. Also, even though sequential adjustments require a large number of multiplications, combining the actual adjustment and blunder detection algorithm into one step offers much greater computational efficiency. Data rejection for a sequential adjustment simply requires deletion of a rejected observation which is much faster than obtaining an new simultaneous solution.

The simultaneous adjustment and blunder detection technique is also very practical. There are many situations where data is added to an adjustment after the adjustment has already been completed and old observations discarded. The sequential algorithm permits the new data to be examined and rejected if required. This procedure may prevent bad data acquired later from ever contaminating a good data set. One may also examine the effect of the observations on the parameter estimates and the covariance matrix of these estimates. Then, there is the possibility for remeasurement to avoid losing a data point. Finally, the sequential adjustment technique can be run on small computers as well as large, and may

provide a blunder detection capability for microcomputers that was previously unavailable.

3.2 Limitations of the Algorithm

The proposed algorithm has two limitations. The first of these is that the residuals must be distributed as a multivariate normal distribution. If this is not the case then the F-test is not valid because the two CHI-SQUARE distributions forming the test statistic are not, in fact, CHI-SQUARE distributions. The normality assumption is not a particularly serious limitation since there are many real world problems about which this assumption is valid. However, one should exercise caution when working with an untested system. The second limitation requires more attention. A sequential adjustment relys on an initial solution. This initial solution is the foundation upon which the entire system will be adjusted and the final results and quality of the blunder detection performance judged. A good initial solution requires enough correct data and initial redundancy to detect small blunders early in the sequential adjustment and prevent a case where acceptable observations become unacceptable as the estimates of the parameters converge. This requirement for a good initial solution requires more discussion.

3.3 The Initial Solution

It is important that the definition of an initial solution be clearly understood. There are two approaches to obtaining an initial solution for a sequential adjustment. One method involves performing a simultaneous adjustment on the first (u+rl) observations. These observations must be selected so that the system is solvable. The rank of the normal equations coefficient matrix must be at least u. The other begins using estimates of the parameters and and the sequential adjustment procedure with the initial observation. The key point is that from the point where the number of observations equals the number of parameters, the two solutions are identical. When (u+rl) observations have been added, both approaches have the same solution. However, initial solution for blunder detection is not merely the starting point for the sequential adjustment. For blunder detection, initial solution refers specifically to the solution upon which the blunder detection algorithm employed in the sequential adjustment will be statistically meaningful. Either method may be used as long as the first (u+rl) observations are blunder free and provide sufficient redundancy to make the F-test meaningful for the observations to be added early in the sequential adjustment.

One critical task of the sequential adjustment is to obtain the (u+rl) blunder free observations for the initial solution. If blunders are present in the initial solution, the entire sequential adjustment is invalid because it proceeds from an incorrect starting point. The F-test of the algorithm is invalid also because the quadratic form in the denominator becomes artificially inflated, permitting even more blunders to be incorporated into the solution as the adjustment proceeds. The problem obviously requires that these observations be taken with particular care but does not pose an insurmountable obstacle for the sequential algorithm to be implimented. One obvious measurement technique would be to make multiple measurements for the observations of the initial solution. As Stefanovic states, multiple measurements do not increase the precision of the measurement, but they do help identify blunders. (Stefanovic, 1981) In this manner, the initial observations can be ' checked out ' and obvious blunders eliminated using nonstatistical methods similar to those of El Hakim (El Hakim & Forstner, 1984) All remaining observations need only be measured once because the algorithm will detect the blunders in this group. The practice of taking multiple measurements is also directly related to the concept of weights. The initial solution should always contain the observations with the highest possible weight. The

weight of the observation reflects certainty, whether through multiple measurements or through more precisely made measurements such as first-order versus second-order survey coordinates. Such an approach is very logical but must sometimes be violated. One must always insure that the initial observations form a well conditioned algebraic system and there should be no singularities in the normal equation coefficient matrix. The sequential adjustment can be simulated to identify these restrictions and the initial observations made accordingly. Finally, one may use the simultaneous adjustment blunder detection algorithms, such as those proposed by Pope and Stefanovic, to search for blunders in the initial solution. The initial solution, though, requires fewer brute searches for the blunders than would a simultaneous adjustment containing all of the observations. The sequential algorithm does not require any more care in measuring than do the simultaneous methods. After the initial solution is obtained, there are fewer measurements required for the sequential adjustment because it is unnecessary to make the multiple measurements for the remaining observations. Therefore, the apparent limitation of requiring good data for the initial solution can be overcome through proper data measurement and point selection methods.

The other key task in the sequential adjustment is to select the proper initial redundancy. This task is important for two reasons. First, the F-test should work as well for data added early in the adjustment as for data added later when the redundancy is higher. If the initial redundancy is too low, then the F-test will be too insensitive. This may mean that observations accepted early in the sequence might have been rejected if added to the soultion later when the redundancy was higher. The test may be particularly insensitive when the first few observations are added and the number of observations being added sequentially (p) is equal to one. The second reason, related to the first,is that if the initial test is too insensitive, blunders may still be allowed into the solution before enough redundancy is achieved to make them statistically detectable.

There are at least two means of determining if the initial redundancy is sufficient. One method is to check if the observations being added after the initial solution produce the same results when added in reverse order. If the results are the same, then the test performed well for observations added early and late in the adjustment. If the results are not the same, then the initial redundancy was not sufficient to insure that the observations added early were properly tested. In this case, one should increase the ini-

tial redundancy. Adding the data in reverse order is a simple empirical test that can be run on different sets of data to establish the desired initial redundancy for a given measurement system. Often, examination of the F-tables for the required level of significance and p will give a good idea of what initial redundancy is required before the F-statistic is small enough to provide a meaningful test. The selection may be subjective or based upon previous experience with a given measurement system.

An alternate approach would be to compute the maximum undetectable blunder. (Stefanovic, 1981) The maximum undetectable blunder could be computed for the initial solution and the final solution for comparison. If they are nearly equal, then the initial solution contained enough initial redundancy. Nearly equal can be either subjective, or the result of propagating the difference between the initial solution maximum undetectable blunder and the final solution maximum undetectable blunder through the solution to determine if the there is a noticeable adverse effect. This is also an empirical test and will probably differ with different measurement systems. Finally, the level of significance also impacts the way the test treats data added early and will be discussed in the next section.

3.4 Type I and Type II statistical Errors

The selection of the level of significance for the F-test in a sequential adjustment is very important because it is also the probability of a Type I error. The Type I error occurs when one rejects the null hypothesis when the null hypothesis is actually true. A type II error occurs when one fails to reject a false null hypothesis. For the statistical test outlined in Section 2.4, the Type I error results in rejecting a good data point while the Type II results in allowing a blunder into the solution. The costs associated with the Type I error include that of remeasurement or requiring a more sophisticated matrix inversion method due to weakened geometry. Type II errors mean that the adjustment contains blunders and is invalid for its intended purpose. Type I and Type II errors have an inverse relationship, meaning that decreasing the level of significance increases the probability of a Type II error. In order to insure that the probability of a Type II error is as small as possible, the level of significance should be as large as possible. A larger level of significance will require that more data be rejected and will require lower initial redundancy for the F-statistic to become small enough to have statistical meaning.

There have been many articles concerning the optimum levels of Type I and Type II errors. Molennaar provides a summary of the subject (Molenaar, 1981). However, there are some comments that should be made concerning these errors and a sequential blunder detection technique. First, one must assume that the goal is to minimize costs. Therefore, any decision made concerning the optimum level of significance should result in the least expensive result. Type II errors usually have the most expensive consequences because blunders adversely effect the final solution. Type I errors, though, may be overcome in several ways. First, the sequential adjustment permits the observation to be remeasured and added later. On line measuring systems can then employ large significance levels because the costs of remeasurement are low. A high significance level is also desired when adding data to an old solution. Too many good data sets have been contaminated after the fact because observations were added later that did not conform to the original data. The high level of significance serves as a tighter filter. Also, the sequential adjustment method can be simulated, and contingencies made for data rejection. Thus, measurements in excess of those required can be made, and a higher level of significance used without having to obtain additional measurements later. Finally, the data does not always have to be rejected. The weights can be altered, but this can be expensive as well.

A key point concerning Type I and Type II errors is that the sequential adjustment provides far greater flexibility for dealing with Type I errors than does the simultaneous adjustment. Such flexibility with Type I errors, in turn, provides better opportunities for preventing far more costly Type II errors. Again, each measurement system should be empirically tested to resolve which level of significance is optimal for the adjustment.

4. Testing the Algorithm

4.1 Digital Sequential Adjustment Blunder Detection

Blunder detection algorithms are most practical when they can be efficiently incorporated into digital computer adjustment programs. The sequential algorithm is particularly easy to program because key quantities for the F-test are computed during the actual adjustment. Appendix F demonstrates that a key portion of the Kalman gain matrix can be used to compute the numerator of the test statistic. Sequential reduction, after data rejection, is also very simple, requiring only several sign changes in the input data. A sequence of critical steps for a computer program for a sequential adjustment with the F-test for blunder detection is in Appendix K. The actual program used to test the algorithm is written in Fortran 77 for an IBM - 370 computer, and the program listing is provided in Appendix K.

4.2 The Data to be Tested

The purpose of this test is to illustrate the flexibility of the algorithm, its programmability, and the types of questions that must be answered when applying the algorithm to a specific measuring system. A straight line is the simplest of models, is easily understood, and does not require linearization. There are two methods of obtaining test data for a straight line. The first method consists of generating a perfect data set and then introducing random error and blunders. The second method requires actual measurements on an actual straight line to be made and the measurements become the observation data. Either method will provide a test data set, however, the second method is used for three reasons. First, the requirement for making measurements is more realistic and does not provide a priori knowledge of a blunder's existence or location. Second, the concept of multiple measurements for the observations to be incorporated into the initial solution can be employed. Finally, the use of a perfect data set with random errors and blunders introduced may not reflect the variability of an actual measuring system and can thus result in an incomplete

understanding of how the algorithm is performing. Any test conducted on a set of actual measurements will simply identify those data points that are statistically inconsistent with the remaining observations without any a priori knowledge of blunders.

Data set one consists of thirty points that lie along a straight line. Tables 4.1 and 4.2 contain the data sets with initial redundancy of 3 and 4 respectively. The first data points for each set have a weight of five because these measurements were made five times. These are the points that will be used in the initial solution. All other points were measured once and have a weight of unity. At the top of each data set are the total number of observations, number of model parameters, the initial redundancy, and number of observations to be added sequentially. The data can be added sequentially in any order desired. The table entries are the observation, the coefficients of the model, the weight of the observation, and the model constant (D(i)) which is zero for this case. For the straight line, the model is

$$Y = A X + B$$

where A is the slope parameter, B is the intercept parameter, Y is the Y-coordinate which is measured, and X is the X-coordinate which is assumed to be error free. The

results will be same as those for an adjustment where both X and Y are treated as observations. The choice of initial redundancy for data set one is subjective. The F-statistic does not appear small enough for observations added sequentially one at a time until the initial solution has a redundancy of three. The object is to select an initial redundancy high enough to enable the algorithm to detect blunders early in the solution. The straight line has two parameters so that an initial redundancy of three requires five observations for the initial solution. The abscissa range is 0 to 7.25 inches, and the ordinate range is 0 to 5.05 inches. Measurements were made with an engineer's scale (60 scale) and converted to decimal inches. The reference variance is not known a priori, but this is not important because the algorithm is not based upon an a priori reference variance. The concept of multiple measurements was used to select data for an initial solution. Finally, the points measured for the initial solution are spread over the entire length to insure that there are no weaknesses in the geometry as observations are added sequentially. Data set two is the same as data set one except the point for X=6.00 is measured four more times and added to the initial solution. Data set two, then, has an initial redundancy of 4.

Table 4.1 - Data Set 1 , Initial Redundancy = 3

n = 30

u = 2

r1 = 3

p = 1

Obs. no.	Y	X	INT	W	D
1	0.672	-1.00	-1.0	5	0.0
2	1.353	-2.00	-1.0	5	0.0
3	2.023	-3.00	-1.0	5	0.0
4	3.375	-5.00	-1.0	5	0.0
5	4.713	-7.00	-1.0	5	0.0
6	0.167	-0.25	-1.0	1	0.0
7	0.333	-0.50	-1.0	1	0.0
8	0.492	-0.75	-1.0	1	0.0
9	0.842	-1.25	-1.0	1	0.0
10	1.008	-1.50	-1.0	1	0.0
11	1.175	-1.75	-1.0	1	0.0
12	1.508	-2.25	-1.0	1	0.0
13	1.683	-2.50	-1.0	1	0.0
14	1.850	-2.75	-1.0	1	0.0
15	2.175	-3.25	-1.0	1	0.0
16	2.342	-3.50	-1.0	1	0.0
17	2.525	-3.75	-1.0	1	0.0
18	2.683	-4.00	-1.0	1	0.0
19	2.842	-4.25	-1.0	1	0.0
20	3.025	-4.50	-1.0	1	0.0
21	3.192	-4.75	-1.0	1	0.0
22	3.533	-5.25	-1.0	1	0.0
23	3.683	-5.50	-1.0	1	0.0
24	3.867	-5.75	-1.0	1	0.0
25	4.025	-6.00	-1.0	1	0.0
26	4.200	-6.25	-1.0	1	0.0
27	4.367	-6.50	-1.0	1	0.0
28	4.550	-6.75	-1.0	1	0.0
29	4.883	-7.25	-1.0	1	0.0
30	5.050	-7.50	-1.0	1	0.0

Table 4.2 - Data Set 2, Initial Redundancy = 4

n = 30

u = 2

rl = 3

p = 1

Obs. No.	Y	X	INT	W	D
1	0.672	-1.00	-1.0	5	0.0
2	1.353	-2.00	-1.0	5	0.0
3	2.023	-3.00	-1.0	5	0.0
4	3.375	-5.00	-1.0	5	0.0
5	4.713	-7.00	-1.0	5	0.0
25	4.041	-6.00	-1.0	5	0.0
6	0.167	-0.25	-1.0	1	0.0
7	0.333	-0.50	-1.0	1	0.0
8	0.492	-0.75	-1.0	1	0.0
9	0.842	-1.25	-1.0	1	0.0
10	1.008	-1.50	-1.0	1	0.0
11	1.175	-1.75	-1.0	1	0.0
12	1.508	-2.25	-1.0	1	0.0
13	1.683	-2.50	-1.0	1	0.0
14	1.850	-2.75	-1.0	1	0.0
15	2.175	-3.25	-1.0	1	0.0
16	2.342	-3.50	-1.0	1	0.0
17	2.525	-3.75	-1.0	1	0.0
18	2.683	-4.00	-1.0	1	0.0
19	2.842	-4.25	-1.0	1	0.0
20	3.025	-4.50	-1.0	1	0.0
21	3.192	-4.75	-1.0	1	0.0
22	3.533	-5.25	-1.0	1	0.0
23	3.683	-5.50	-1.0	1	0.0
24	3.867	-5.75	-1.0	1	0.0
26	4.200	-6.25	-1.0	1	0.0
27	4.367	-6.50	-1.0	1	0.0
28	4.550	-6.75	-1.0	1	0.0
29	4.883	-7.25	-1.0	1	0.0
30	5.050	-7.50	-1.0	1	0.0

Note: Observation 25 was remeasured 4 more times and added to the initial solution data to change Data Set 1 into Data Set 2.

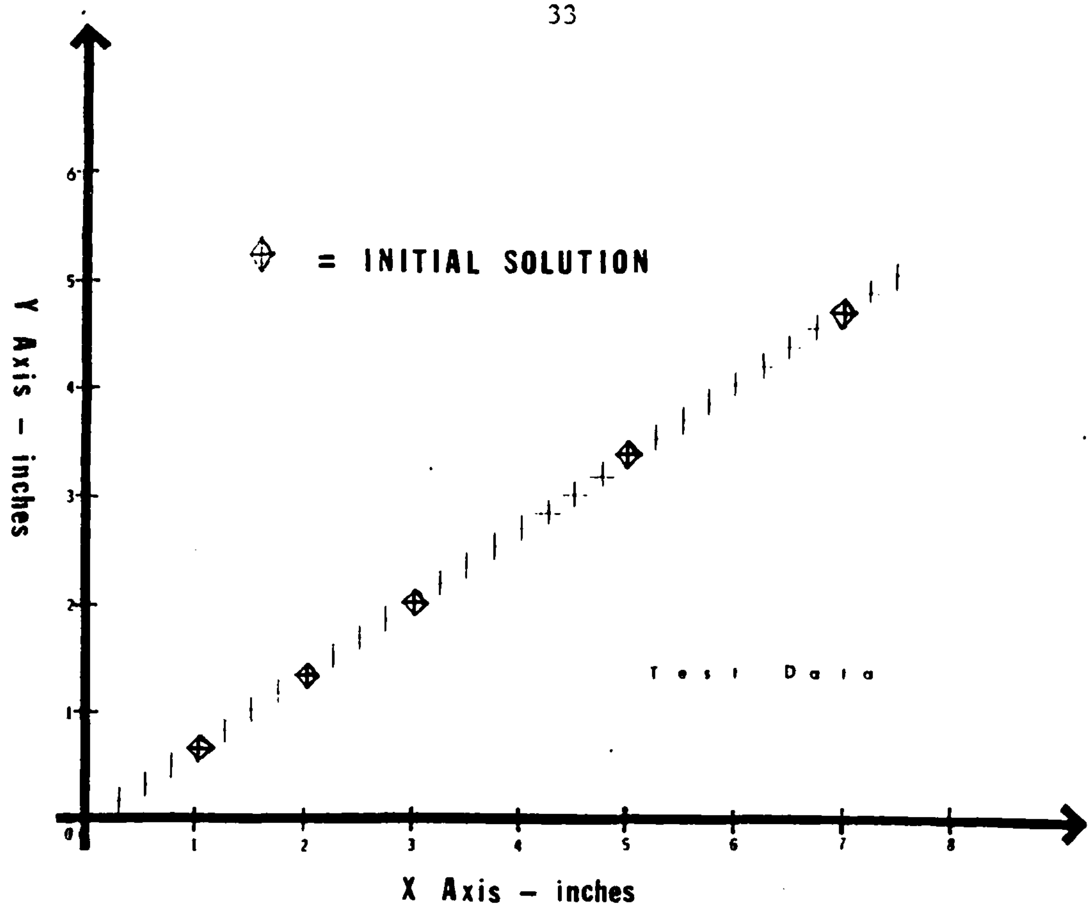


FIGURE 4.1 TEST DATA PLOT

4.3 Experiment Design

Six tests are required to illustrate how the blunder detection algorithm works. First, three tests are run on data set one to determine the effect of altering the level of significance while initial redundancy and the order in which the data is added sequentially are kept constant. Tests 1,2, and 3 should show that more data is rejected as the level of significance becomes larger. Then, an additional test should be run on data set one at the same level of significance as Test 3, but with the observations being added sequentially in reverse order. Test 4 should show that if the initial solution does not contain enough redundancy, or is otherwise incorrect, the same data points will not be rejected when the order in which data is added is changed. The last two tests are run on data set two in the same manner as Tests 3 and 4 are run on data set one. However, data set two has a greater initial redundancy. Tests 5 and 6 may show that the same data points are rejected if the order in which the data is added is reversed. The increase in initial redundancy strengthens the initial solution and enables it to detect blunders early in the adjustment. Finally, the same data points should consistently be rejected for this group of tests. A lack of commonality of data rejected may indicate that the algorithm is

Table 4.3 - Tests To Be Run

Test No.	Data Set	rl	Alpha	Test Effect Of:
1	1	3	.10	Level of Significance
2	1	3	.01	Level of Significance
3	1	3	.05	Level of Significance
4	1	3	.05	Adding Data in Reverse order
5	2	4	.05	Adding Extra Point to Initial Solution
6	2	4	.05	Adding Data in Reverse Order

not performing as it should. Table 4.3 lists each test and its purpose. Test results should be considered by comparing the results of test groupings. Tests 1,2, and 3 are the group that illustrate the effects of varying level of significance. Tests 3 and 4 should show that the order in which data is added can effect the results when the initial redundancy is too low. And, tests 5 and 6 should be compared with the results of tests 3 and 4 to show that an increase in initial redundancy may correct the problem and insure that order has no effect on which data points are rejected.

4.4 Test Results

The six tests demonstrate the use of the sequential adjustment algorithm proposed in this paper. Tables 4.4 and 4.5 summarize the results. Table 4.4 shows the number of observations rejected for each test as the level of significance, initial redundancy, or order in which data are added sequentially are varied. Tests 1,2, and 3, conducted on data set one, demonstrate how a higher level of significance causes the algorithm to reject more data. Tests 3 and 4 show that when level of significance and initial redundancy are constant, the order in which observations are added sequentially can alter which data points are rejected. This is an indica-

tion that the initial redundancy may be too low. Tests 5 and 6 reinforce this conclusion because by adding an extra point of high weight to the initial solution, the same observations are rejected regardless of whether the data is added in the listed order or reverse order. Tests 1 and 4 rejected no observations.

Table 4.5 shows which points were rejected by each test. As expected, tests 2,3,5, and 6 rejected common points. The observations corresponding to X=4.25 and X=5.50 were rejected each time and were even borderline points for test 4. This commonality is important because it indicates that these two points really are blunders and that the algorithm is performing as it should. Test 2 shows five points rejected. This larger number of points rejected is also expected because test 2 had the highest level of significance.

These six tests, run on data sampled from a straight line, are representative of those which would be required for any measurement system if the proposed blunder detection technique is to be used. The tests illustrate the need to carefully plan measurements, the requirement for a strong initial solution, and the effects of varying the level of significance. The test results also confirm that the algorithm works, at least for the straight line model. The same

Table 4.4 - Test Results For Straight Line Data

Test	rl	Data Set	Alpha	No. Obs. Rej.
1	3	1	.01	None
2	3	1	.10	5
3	3	1	.05	2
4	3	1	.05	None
5	4	2	.05	2
6	4	2	.05	2

Note: Tests 4 and 6 were conducted by adding the sequential data in reverse order.

Table 4.5 - Data Rejected For Each Test

Test	X Coord. Rej.	Obs. No.	FSTAT1	FSTAT2
2	3.25	15	3.67	3.18
2	3.50	16	4.45	3.18
2	4.25	19	7.26	3.10
2	5.50	23	7.93	3.03
2	6.00	25	4.48	3.01
3	4.25	19	5.06	4.49
3	5.50	23	5.82	4.38
5	4.25	19	5.32	4.45
5	5.50	23	6.20	4.35
6	4.25	19	7.36	4.67
6	5.50	23	7.00	4.96

Data rejected when FSTAT1 > FSTAT2.

FSTAT1 = Test Statistic

FSTAT2 = $F(\alpha, p, r1)$

For test 4, the F values for X = 4.25 , 5.5:

X = 4.25 FSTAT1= 3.91 FSTAT2 = 4.60

X = 5.50 FSTAT1 = 4.72 FSTAT2 = 4.96

data points are consistently rejected, and the algorithm behaved as expected. Once the initial redundancy was increased, the same observations were rejected when they were added sequentially in reverse order. These tests are not intended to be rigorous for all measurement systems, but they do illustrate the use of the algorithm, its programmability, and the types of questions that must be answered when applying the algorithm to a specific measuring system.

5. Concluding Remarks

5.1 Summary of the Algorithm Proposed

The stated objective of this thesis was to develop an algorithm for detecting blunders in observed data using a sequential least squares adjustment technique. The algorithm developed consists of an F-test that compares two key quadratic forms. The test statistic,

$$T.S. = \frac{[(V_2^t \bar{Q}_{22}^{-1} V_2)/p]}{[(\hat{V}_1^t W_1 \hat{V}_1)/r1]} \qquad 5.1$$

is easily computed for each step of the sequential adjustment. The 'p' observations being added sequentially are rejected if the T.S. $> F(\alpha, p, r1)$. Tests on measured data sampled along a straight line model indicate that the algorithm works well on that simple model as long as the initial solution contains enough redundant observations of high weight to insure the test is valid for the observations added early in the adjustment. The requirements for the algorithm have been met as follows:

1. This algorithm is suitable for detection of blunders of small and large magnitude. Large magnitude refers

to those blunders large enough to be detected using non-statistical methods (El-Hakim and Zeiman, 1984). The sequential technique does not require, however, a special step to find these large blunders because they will be rejected before they contaminate the solution and inflate the covariance matrix of the parameter estimates. However, the initial solution must contain enough redundant observations of high weight to insure that the algorithm works well early in the adjustment for blunders of small magnitude.

2. Blunder location is given immediately.
3. The algorithm is computationally efficient. It combines the blunder detection, data rejection, and the actual adjustment into one step. The simutaneous adjustment requires an adjustment followed by a search for blunders, and another simultaneous adjustment if data are rejected as blunders.
4. Tolerances may be determined by means of confidence intervals using the a posteriori reference variance that is updated after each observation is added in sequence. Also, the maximum undetectable blunder may be calculated for the initial and final solutions to insure that the algorithm works well early in the sequential adjustment.
5. A priori knowledge is not required concerning number

or location of blunders in the observations that are added sequentially. However,the presence of very large numbers of blunders relative to the number of parameters indicates a problem with either model or data transmission. Also, there must be adequate blunder free data to form a good initial solution.

6. The statistical measures of confidence for Type I and Type II errors are incorporated in the algorithm and may be varied to meet the requirements of the user. Type I errors are controlled directly through selecting a specific level of significance. Using a higher probability of a Type I error permits one to indirectly control the probability of a more expensive Type II error.

As far as the challenges listed in Section 1.1 are concerned, they can be largely overcome. The inversion of large matrices may be avoided using a sequential technique beginning with the first observation and can permit the algorithm to be used on personal computers. The fact that the true residuals are not available for each step of the adjustment does not prohibit the detection of blunders. As demonstrated in Appendix E, the residuals may be expressed in terms of residual estimates at each step of the sequential adjustment, allowing the use of the key quadratic forms in equations 2.6 and 2.7. Finally,

the difficulty in establishing the initial redundancy and initial solution can be solved rather easily for the simple straight line tested. That test illustrated that it is desireable to establish an initial solution strong enough to allow the algorithm to work well early in the adjustment. A strong initial solution will also preclude accepting observations early in the sequential adjustment that would be rejected as blunders once the parameter estimates begin to converge. There are several questions concerning the initial solution that still require further study. These questions will be briefly discussed in Section 5.3.

5.2 Applications for Sequential Blunder Detection

It is important to note that the sequential adjustment blunder detection algorithm may be used in place of any other least squares adjustment method. Sequential adjustments produce results identical to simultaneous adjustments. The key advantage for the sequential approach is that it provides much greater flexibility than do simultaneous methods. In cases where there are already solutions, particularly those where observations have been discarded, new data can be tested to insure that the old solution does not become conta-

minated if the new data does contain blunders. The sequential approach also permits blunder detection for those systems that rely on a sequential adjustment for the solution. Orbit determination is one such problem. There are also many opportunities to employ a sequential technique to on line measurement systems. Finally, this algorithm is statistically based and does not rely on subjective judgement that may or may not be correct. Artificial techniques such as a three-sigma cutoff or estimating $\bar{Q}$ with Q of the observations are weak, have no statistical basis, and should not be used. The computer programs for sequential adjustment are no more complex than elimination methods and offer several computational efficiencies for blunder detection. The 'artificial methods' appeal due to their computational efficiency is no longer valid, particularly when one considers that their computational efficiency is seriously degraded if the wrong observations are discarded.

5.3 Topics for Further Study

Most of the difficulties with implementing the sequential adjustment algorithm involve achieving the initial solution. While there are some empirical tests that may be used, meth-

ods such as running the adjustment in reverse order are somewhat wasteful and should be avoided if possible. However, the requirement for a blunder free initial solution cannot be taken too lightly and poses the following questions:

1. Is there a way to determine initial redundancy without resorting to empirical testing?
2. Is there an optimum configuration for selecting the data points for the initial solution for various models?
3. How does one insure that the observations in the initial solution are blunder free?
4. Can this algoithm be extended to permit checking all of the observations sequentially, even those in the initial solution, thus permitting a true on-line check for applications in real time mensuration.
5. Can this algorithm be extended to the case where parameters as well as observations are being added?

Finally, there are several questions raised that are not directly related to the initial solution. First, the algorithm was tested on a simple, linear model. Tests should be conducted to determine if there is a requirement to alter the test statistic for cases when nonlinear adjustment problems require more than one iteration to converge. Also, the ever

present question of weights persists. In addition to simply trying to identify weights, is there an optimal way to sequentially add observations having higher weights? The answer would seem to be that the more heavily weighted observations should be added as early as possible since weight denotes certainty in the measurement. It is desirable to strengthen the solution early to catch early blunders more efficiently. However, there are questions of geometry, dispersion or range of the data, and so on, that may require data to be added in a different manner. Many of these questions will require testing actual measurement systems to test the limitations and capabilities of the sequential adjustment blunder detection algorithm.

REFERENCES

Amer, F. "Theoretical Reliability of Elementary Photogrammetric Procedures.", ITC Journal, 1981-3, (1981), 278-301.

Amer, F. "Aerial Triangulation Accuracy Improvement and Quality Control.", ITC Journal, 1980-4, (1980), 582-605.

Baarda, W. "Statistical Concepts in Geodesy", Netherlands Geodetic Commission, DELFT, Publications in Geodesy, 1967.

Barnet, V., and Lewis, T. Outliers in Statistical Data. New York: John Wiley and Sons, 1978.

Bjerhammer, Arne. Theory of Errors and Generalized Matrix Inverses. Amsterdam: Elsvier Scientific Publishing Company, 1967.

Bouloucos, T. "Error Detection and Reliability Studies in Analyitically Formed Strips.", ITC Journal, 1981-3, (1981), 60-69.

El-Hakim, S. F., and Zieman. "A Step by Step Strategy for Gross Error Detection", Photogrammetric Engineering and Remote Sensing, 50(6), (1984), 713-718.

Mikhail, M. Observations and Least Squares. Washington D.C.: University Press of America, 1967.

Molenaar, M. "Error Detection in Plainimetric Block Adjustment" , ISP Commission III, Presented Paper, 13th Congress of the ISP, Helsinki, 1976.

Molenaar, M. "Some Problem Areas in the Quality Control for Aerotriangulation", ITC Journal, 1981-3, (1981), 253-265.

Molenaar, M. "Risk Minimization for Error Detection Procedures in Photogrammetry", ITC Journal, 1981-2 (1981), 185-201.

NOAA. On Data Snooping and Multiple Outlier Testing. By J. Kok. NOAA Technical Report NOS NGS 30. Rockville, Maryland: GPO, 1984.

NOAA. <u>The Statistics of Residuals and Detection of Outliers</u>. By A. Pope. NOAA Technical Report NOS65NGS1. Rockville, Maryland: GPO, 1976.

Stefanovic, P. "Blunders and Least Squares" , <u>ITC Journal</u>, 1978-1, (1978), 122-155.

Stefanovic, P."Pitfalls in Blunder Detection", <u>ITC Journal</u>, 1981-1, (1981), 81-97.

Appendix A : Abbreviations and Notations.

X - the vector of estimates of the parameters.

L - the vector of observations.

e - the vector of observational errors.

Y - the vector of expectations of the observations.

V - the vector of residuals.

A - the coefficient matrix of the observations.

B - the coefficient matrix of the parameters/unknowns.

n - the total number of observations in a least squares solution.

u - the total number of parameters/unknowns.

r - the redundant number of observations (r=n-u).

b - the maximum allowable number of blunders in a data set.

W - the weight matrix of the observations.

Q - the cofactor matrix of the observations (equals the inverse of W).

$\bar{Q}$ - the cofactor matrix of the residuals.

c - the number of condition equations in a least squares solution.

Ho - the symbol for the null hypothesis in statistics.

Ha - the symbol for the alternate hypothesis.

α - the level of significance of a statistical test.

t - is the symbol for transpose of an individual matrix. B^t is B(transpose) for example.

J - the symbol refering to the Jacobian of a matrix.

Appendix B : Sequential Adjustments and Notations.

1. The general least squares model may be written in matrix form as

$$A(L+V)+BX=D \tag{B1}$$

or

$$AV+BX=F \tag{B2}$$

where

A has dimensions c by n,

D,F have dimensions c by 1,

L,V have dimensions n by 1,

B has dimensions c by u,

X has dimensions u by 1.

2. However, the technique of adjustment by indirect observations produces results identical to those above and is denoted by

$$A(L+V)+BX=D \tag{B3}$$

or

$$(L+V)+BX=D \tag{B4}$$

or

$$V+BX=F \tag{B5}$$

where

A is n by n identity matrix,

B has dimensions n by u,

X has dimensions u by 1,

F,L,V have dimensions n by 1.

3. Equation (B3) may be partitioned to denote the contribution to the total solution by components of a sequential solution. Subscript (1) refers to the portion of the model associated with observations already included in the solution. Subscript (2) refers to the ' p ' new observations being added sequentially to the solution. In matrix notation,

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} X_2 = \begin{bmatrix} F_1 \\ F_2 \end{bmatrix} \tag{B6}$$

where the matrices have the following dimensions:

V_1 is n by 1 $\qquad$ V_2 is p by 1

B_1 is n by u $\qquad$ V_2 is p by u

F_1 is n by 1 $\qquad$ F_2 is p by 1

X_2 is u by 1

and,

X_1 is the vector of parameter estimates obtained using only the observations L_1.

X_2 is the vector of parameter estimates obtained using both L_1 and L_2.

4. The solution for X_2 in equation (B6) is

$$X_2 = (N_2)^{-1} [T_1 + T_2] \tag{B7}$$

where

$$N_2 = [B_1^t W_1 B_1 + B_2^t W_2 B_2] \tag{B8}$$

$$[T_1 + T_2] = [B_1^t W_1 F_1 + B_2^t W_2 F_2] \tag{B9}$$

$$W = \begin{bmatrix} W_1 & 0 \\ 0 & W_2 \end{bmatrix}$$

where

W_1 has dimensions n by n,

W_2 has dimensions p by p.

5. The residual vector obtained by considering only L_1 is

$$\hat{V}_1 + B_1 X_1 = F_1 \tag{B10}$$

where

$$X_1 = [B_1^t W_1 B_1]^{-1} [B_1^t W_1 F_1] \tag{B11}$$

$$N_1 = [B_1^t W_1 B_1]$$

6. Notations for the residuals

a. V_1, V_2 refer to vectors of residuals from the solution of B6. L_1 and L_2 are included in the solution of X_2.

b. $\hat{V}_1$ refers to the residuals obtained from the solution

of B10 when only L_1 is included in the solution of X_1.

c. $\hat{V}_2$ refers to the discrepancies equal to $F_2 - B_2X_1$.

$\hat{V}_2$ is one estimate of V_2 and is termed the discrepancy vector.

d. $V_1, V_2, \hat{V}_1$ are all residuals. $\hat{V}_2$ is a discrepancy vector.

7. Formulas for residuals and discrepancies follow:

$$V_1 = F_1 - B_1X_2 \tag{B12}$$

$$V_2 = F_2 - B_2X_2 \tag{B13}$$

$$\hat{V}_1 = F_1 - B_1X_1 \tag{B14}$$

$$\hat{V}_2 = F_2 - B_2X_1 \tag{B15}$$

Appendix C : Sequential Least Squares Adjustments.

1. From equations (B8) and (B11),

$$N_2 = [B_1^t W_1 B_1 + B_2^t W_2 B_2] = [N_1 + B_2^t W_2 B_2] \tag{C1}$$

2. The inverse of N_2 can easily be found using the well known matrix relationship: (Mikhail, 1976)

If X=[Y +UZV] , then (C2)

$$X^{-1} = Y^{-1} - Y^{-1} U [Z^{-1} + V Y^{-1} U]^{-1} V Y^{-1} \tag{C3}$$

3. Step 2 gives the following result for N_2:

$$N_2^{-1} = N_1^{-1} - N_1^{-1} B_2^t [Q_2 + B_2 N_1^{-1} B_2^t]^{-1} B_2 N_1^{-1} \tag{C4}$$

4. Also, letting $C_1 = N_1^{-1}$ and $C_2 = N_2^{-1}$,

$$X_2 = C_2 T_2 = C_2 T_1 + C_2 T'_2$$

$$C_2 T_1 = C_1 T_1 - C_1 B_2^t [B_2 C_1 B_2^t + Q_2]^{-1} B_2 C_1 T_1$$

$$C_2 T'_2 = C_1 [B_2^t W_2 F_2] - C_1 B_2^t [B_2 C_1 B_2^t + Q_2]^{-1} B_2 C_1 B_2^t W_2 F_2$$

$$= C_1 B_2^t W_2 F_2 - C_1 B_2^t W_2 F_2 + C_1 B_2^t [B_2 C_1 B_2^t + Q_2] Q_2 W_2 F_2$$

$$= C_1 B_2^t [B_2 C_1 B_2^t + Q_2]^{-1} F_2$$

so,

$$X_2 = C_{21} T_1 + C_{22} T'_2$$

$$= C_1 T_1 + C_1 B_2^t [B_2 C_1 B_2^t + Q_2]^{-1} [F_2 - B_2 C_1 T_1]$$

5. From step 4,

$$X_2 = X_1 + C_1 B_2^t [B_2 C_1 B_2^t + Q_2] [F_2 - B_2 X_1]^{-1} \tag{C5}$$

6. The quantity $C_1 B_2^t [B_2 C_1 B_2^t + Q_2]^{-1}$ is known as the Kalman gain matrix and will be designated K.

7. Therefore,

$$X_2 = X_1 + K [F_2 - B_2 X_1] \tag{C6}$$

and

$$N_2^{-1} = N_1^{-1} - K B_2 N_1^{-1} \tag{C7}$$

8. These equations can be generalized to a form where (i-1) corresponds to subscript (1) and (i) corresponds to the subscript (2). However the subscripts 1,2 simplify the notation and will be used throughout this paper. Any conclusions made concerning examples using this notation can easily be extended to the general form.

9. One may also show that observations may be easily deleted simply by changing the signs of the key matrix W_2. (Mikhail, 1976)

Appendix D : The Cofactor Matrix of the Residuals.

1. The cofactor matrix of the residuals may be expressed in matrix notation (using the method of indirect observations) as

$$\bar{Q} = [\ Q - B\ [N_2]^{-1}\ B^t\ ] \tag{D1}$$

where Q is block diagonal.

2. $\bar{Q}$ may be partitioned and calculated as follows:

$$\bar{Q} = \begin{bmatrix} \bar{Q}_{11} & \bar{Q}_{12} \\ \bar{Q}_{21} & \bar{Q}_{22} \end{bmatrix} = \begin{bmatrix} Q_1 & 0 \\ 0 & Q_2 \end{bmatrix} - \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} [N_2]^{-1} \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}^t \tag{D2}$$

where

$\bar{Q}_{11}$ is the cofactor matrix for the residuals V_1.

$\bar{Q}_{22}$ is the cofactor matrix for the residuals V_2.

$\bar{Q}_{12} = \bar{Q}_{21}$ = the cross-cofactor matrix for V_1 and V_2.

$Q_1 = W_1^{-1}$ and $Q_2 = W_2^{-1}$.

3. Step 2 shows that the cofactor matrix for the residuals of the observations being added sequentially can be calculated at each step since the inverse of N_2 is calculated each time.

$$\bar{Q}_{22} = [\ Q_2 - B_2\ [N_2]^{-1}\ B_2^t\ ] \tag{D3}$$

Appendix E : Discrepancies in Terms of Residuals.

1. Let $V + BX = F$ be partitioned as

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} X = \begin{bmatrix} F_1 \\ F_2 \end{bmatrix} \qquad \text{(E1)}$$

2. Now, complete a sequential reduction by eliminating the observations L_2 that correspond to V_2, B_2, and F_2.

3. Let $N_2^{-1} = [\,B_1^t W_1 B_1 + B_2^t W_2 B_2\,]^{-1}$ and let $\bar{N}^{-1}$ be the inverse of the new normal equation coefficient matrix after the observations L_2 are deleted.

4. Using the results of Appendix C, and calling the new vector of parameter estimates X',

$$X_1 = X - N_2 B_2^t\,[Q_2 - B_2 N_2^{-1} B_2^t\,]^{-1}\,[F_2 - B_2 X_2] \qquad \text{(E2)}$$

5. From (B6),

$$X = X_2.$$

6. Also,

$$\bar{N}^{-1} = [\,B_1^t W_1 B_1 + B_2^t W_2 B_2 - B_2^t W_2 B_2\,]^{-1} = N_1^{-1}.$$

7. $\hat{V}_1 = F_1 - B_1 X_1$ since only observations L_1 are part of the solution. (see Appendix B)

8. Combining (E2), (D3), and step 7 above,

$$\hat{V}_1 = F_1 - B_1 X_2 + [B_1 N_2^{-1} B_2^t] \bar{Q}_{22}^{-1} [F_2 - B_2 X_2] \qquad (E3)$$

9. From (B12) and (B13),

$$F_1 - B_1 X_2 = V_1 \quad \text{and} \quad F_2 - B_2 X_2 = V_2$$

10. From (D2) $B_1 N_2^{-1} B_2^t = -\bar{Q}_{12}$.

11. Therefore,

$$\hat{V}_1 = V_1 - \bar{Q}_{12} \bar{Q}_{22}^{-1} V_2 \qquad (E4)$$

where

$\hat{V}_1$ refers to the residuals for observations L_1 after the effect of observations L_2 have been removed from the solution.

12. Taking this a step further, a relationship between the residuals V_2 and discrepancies $\hat{V}_2$ can be developed.

$$\hat{V}_2 = F_2 - B_2 X' = F_2 - B_2 [X_2 - N_2^{-1} B_2 [Q_2 - B_2 N_2^{-1} B_2^t]^{-1} [F_2 - B_2 X_2]]$$

$$= F_2 - B_2 X_2 + B_2 N_2^{-1} B_2^t \bar{Q}_{22}^{-1} [F_2 - B_2 X_2]$$

$$= [I + B_2 N_2^{-1} B_2^t \bar{Q}_{22}^{-1}] V_2$$

13. And,

$$I + B_2 N_2^{-1} B_2^t \bar{Q}_{22}^{-1} = I + (Q_2 - \bar{Q}_{22}) \bar{Q}_{22}^{-1}$$

$$= I + Q_2 \bar{Q}_{22}^{-1} - I = Q_2 \bar{Q}_{22}^{-1}$$

14. Therefore,

$$\hat{V}_2 = Q_2 \bar{Q}_{22}^{-1} V_2 \quad \text{or} \quad V_2 = \bar{Q}_{22} W_2 \hat{V}_2 \tag{E5}$$

Appendix F : A Key Equality

1. Show that

$$V_2^t \bar{Q}_{22}^{-1} V_2 = \hat{V}_2 [Q_2 + B_2 N_1^{-1} B_2^t]^{-1} \hat{V}_2$$

where

V_2 and $\hat{V}_2$ are defined as in (B13) and B(15).

2. From Appendix E,

$$\hat{V}_2 = Q_2 \bar{Q}_{22}^{-1} V_2 .$$

3. Or,

$$V_2 = \bar{Q}_{22} W_2 \hat{V}_2 .$$

4. Substituting,

$$V_2^t \bar{Q}_{22}^{-1} V_2 = \hat{V}_2^t W_2 \bar{Q}_{22} \bar{Q}_{22}^{-1} \bar{Q}_{22} W_2 \hat{V}_2 \qquad \text{(F1)}$$

since W and $\bar{Q}$ are symmetric matrices.

5. The left side of equation (F1) = $\hat{V}_2^t W_2 \bar{Q}_{22} W_2 \hat{V}_2$ (F2)

6. From (F2) and (D3),

$$W_2 \bar{Q}_{22} W_2 = W_2 (Q_2 - B_2 N_2 B_2^t) W_2$$

$$=[I - W_2 B_2 N_2^{-1} B_2^t] W_2$$

$$=[W_2 - W_2 B_2 N_2^{-1} B_2^t W_2]$$

7. Using the same equation used in (C2), one can show that

$$[Q_2 + B_2 N_1^{-1} B_2^t]^{-1} = W_2 - W_2 B_2 N_2^{-1} B_2^t W_2 .$$

8. Therefore,

$$V_2^t \bar{Q}_{22}^{-1} V_2 = \hat{V}_2^t [Q_2 + B_2 N_1^{-1} B_2^t]^{-1} \hat{V}_2 \qquad \text{(F3)}$$

Appendix G : The Total Sum of Weighted Squared Residuals.

1. The purpose of this appendix is to show the following:

$$V^t W V = \hat{V}_1^t W_1 \hat{V}_1 + V_2^t \bar{Q}_{22}^{-1} V_2 \qquad \text{(G1)}$$

2. The left of (G1) can be partitioned so that

$$V^t W V = \begin{bmatrix} V_1 \\ V_2 \end{bmatrix}^t \begin{bmatrix} W_1 & 0 \\ 0 & W_2 \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = V_1^t W_1 V_1 + V_2^t W_2 V_2 . \qquad \text{(G2)}$$

3. From Appendix E,

$$V_1 = \hat{V}_1 + \bar{Q}_{12} \bar{Q}_{22}^{-1} V_2$$

$$V_2 = \bar{Q}_{22} W_2 \hat{V}_2$$

$$V_1 = \hat{V}_1 + \bar{Q}_{12} W_2 \hat{V}_2$$

4. So,

$$V_1^t W_1 V_1 = [\hat{V}_1 + \bar{Q}_{12} W_2 \hat{V}_2]^t W_1 [\hat{V}_1 + \bar{Q}_{12} W_2 \hat{V}_2] \qquad \text{(G3)}$$

$$= \hat{V}_1^t W_1 \hat{V}_1 + \hat{V}_1^t W_1 \bar{Q}_{12} W_2 \hat{V}_2 + \hat{V}_2^t W_2 \bar{Q}_{21} W_1 \hat{V}_1$$

$$+ \hat{V}_2^t W_2 \bar{Q}_{21} W_1 \bar{Q}_{12} W_2 \hat{V}_2$$

5. However,

$$B_1^t W_1 \hat{V}_1 = \bar{Q}_{21} W_1 \hat{V}_1 = 0 .$$

6. Therefore,

$$V_1^t W_1 V_1 = \hat{V}_1^t W_1 \hat{V}_1 + \hat{V}_2^t W_2 \bar{Q}_{21} W_1 \bar{Q}_{12} W_2 \hat{V}_2 \tag{G4}$$

7. Similarly,

$$V_2^t W_2 V_2 = \hat{V}_2^t W_2 \bar{Q}_{22} W_2 \bar{Q}_{22} W_2 \hat{V}_2 \tag{G5}$$

8. Combining (G2), (G4), and (G5),

$$V^t W V = \hat{V}_1^t W_1 \hat{V}_1 + \hat{V}_2^t W_2 \bar{Q}_{21} W_1 \bar{Q}_{12} W_2 \hat{V}_2 + \hat{V}_2^t W_2 \bar{Q}_{22} W_2 \bar{Q}_{22} W_2 \hat{V}_2 \tag{G6}$$

9. From (G1) and (G6), it is only necessary to show that

$$G = \hat{V}_2^t W_2 \bar{Q}_{21} W_1 \bar{Q}_{12} W_2 \hat{V}_2 + \hat{V}_2^t W_2 \bar{Q}_{22} W_2 \bar{Q}_{22} W_2 \hat{V}_2 = V_2^t \bar{Q}_{22}^{-1} V_2 . \tag{G7}$$

10. Using the Distributive Law for matrix multiplication,

$$G = \hat{V}_2^t [W_2 \bar{Q}_{21} W_1 \bar{Q}_{12} W_2 + W_2 \bar{Q}_{22} W_2 \bar{Q}_{22} W_2] \hat{V}_2 .$$

11. Then,

$$W_2 \bar{Q}_{21} W_1 \bar{Q}_{12} W_2 = W_2 B_2 N_2^{-1} N_1 N_2^{-1} B_2^t W_2 \tag{G8}$$

12. And,

$$W_2 \bar{Q}_{22} W_2 \bar{Q}_{22} W_2 = W_2 [Q_2 - B_2 N_2^{-1} B_2^t] W_2 [Q_2 - B_2 N_2^{-1} B_2^t] W_2$$

$$= [W_2 - W_2 B_2 N_2^{-1} B_2^t W_2] [I - B_2 N_2^{-1} B_2 W_2]$$

$$= [W_2 - W_2 B_2 N_2^{-1} B_2^t W_2] - [W_2 B_2 N_2^{-1} B_2^t W_2]$$

$$+ [W_2 B_2 N_2^{-1} B_2^t W_2 B_2 N_2^{-1} B_2^t W_2] \qquad \text{(G9)}$$

13. Combining (G8) and (G9) then gives

$$G = \hat{V}_2^t [W_2 B_2 N_2^{-1} N_1 N_2^{-1} B_2^t W_2 + W_2 B_2 N_2^{-1} B_2^t W_2 B_2 N_2^{-1} B_2^t W_2$$

$$- W_2 B_2 N_2^{-1} B_2^t W_2 + [W_2 - W_2 B_2 N_2^{-1} B_2^t W_2]] \hat{V}_2$$

$$= \hat{V}_2^t [W_2 B_2 N_2^{-1} [N_1 + B_2^t W_2 B_2] N_2^{-1} B_2^t W_2$$

$$- W_2 B_2 N_2^{-1} B_2^t W_2 + [W_2 - W_2 B_2 N_2^{-1} B_2^t W_2]] \hat{V}_2$$

$$= \hat{V}_2^t [W_2 - W_2 B_2 N_2^{-1} B_2^t W_2] \hat{V}_2 \qquad \text{(G10)}$$

14. Using (C2) and (C3) it can be shown that

$$[W_2 - W_2 B_2 N_2^{-1} B_2^t W_2] = [Q_2 + B_2 N_1^{-1} B_2^t]^{-1}$$

$$G = \hat{V}_2^t [Q_2 + B_2 N_1^{-1} B_2^t]^{-1} \hat{V}_2$$

15. Using the results, then, of appendix F

$$
\begin{aligned}
V^t W V &= \hat{V}_1^t W_1 \hat{V}_1 + G \\
&= \hat{V}_1^t W_1 \hat{V}_1 + \hat{V}_2^t [Q_2 + B_2 N_1^{-1} B_2^t] \hat{V}_2 \\
&= \hat{V}_1^t W_1 \hat{V}_1 + V_2^t \bar{Q}_{22}^{-1} V_2 \qquad \text{(G11)}
\end{aligned}
$$

which proves that (G1) is indeed true.

Appendix H : Independent Quadratic Forms.

1. The purpose of this appendix is to show that

$$\hat{V}_1^t W_1 \hat{V}_1 \text{ and } V_2^t \bar{Q}_{22}^{-1} V_2 \text{ are independent.}$$

2. The first step in demonstrating their independence is to demonstrate that $\hat{V}_1$ and V_2 are independent.

3. Returning to the basic principles of error propagation,

$$Q_{XY} = J_{XZ} Q_{ZZ} J_{YZ}^t \tag{H1}$$

4. For this problem,

$$\hat{V}_1 = F_1 - B_1 X_1 = F_1 - B_1 N_1^{-1} [B_1^t W_1 F_1] \tag{H2}$$

$$V_2 = F_2 - B_2 X_2 = F_2 - B_2 N_2^{-1} [B_1^t W_1 F_1 + B_2^t W_2 F_2] \tag{H3}$$

5. The only common variable is F_1 and the crosscofactor matrix becomes

$$Q_{\hat{V}_1 V_2} = J_{\hat{V}_1 F_1} Q_{F_1 F_1} J_{V_2 F_1}^t \;. \tag{H4}$$

6. For indirect observations (observation equations),

$$Q_{F_1 F_1}^{-1} = W_1$$

7. $$J_{\hat{V}_1 F_1} = I - B_1 N_1^{-1} B_1^t W_1 \qquad \text{(H5)}$$

$$J_{V_2 F_1} = B_2 N_2^{-1} B_1^t W_1 \qquad \text{(H6)}$$

8. Substituting,

$$J_{\hat{V}_1 F_1} Q_{F_1 F_1} J^t_{V_2 F_1}$$

$$= [I - B_1 N_1^{-1} B_1^t W_1] [W_1^{-1}] [B_2 N_2^{-1} B_1^t W_1]^t$$

$$= [W_1^{-1} - B_1 N_1^{-1} B_1^t] [W_1 B_1 N_2^{-1} B_2^t]$$

$$= [B_1 N_2^{-1} B_2^t] - [B_1 N_1^{-1} B_1^t W_1 B_1 N_2^{-1} B_2^t]$$

$$= [B_1 N_2^{-1} B_2^t] - [B_1 N_2^{-1} B_2^t] = [0] \qquad \text{(H7)}$$

9. Therefore, $\hat{V}_1$ and V_2 are independent and their respective quadratic forms are also independent because W_1 and $\bar{Q}_{22}$ are matrices of constants.

10. If the assumption is made that the observational errors are normally distributed with a mean of 0 and a variance of σ_0^2, then

$$\hat{V}_1 \sim N\left(0, \Sigma_{\hat{V}_1 \hat{V}_1}\right) \qquad \text{(H8)}$$

$$V_2 \sim N(0, \Sigma_{V_2 V_2}) \tag{H9}$$

11. This further implies that:

$$\hat{V}_1^t W_1 \hat{V}_1 \sim \text{CHI-SQUARE}(r) \times \sigma_0^2 \tag{H10}$$

$$V_2^t \bar{Q}_{22}^{-1} V_2 \sim \text{CHI-SQUARE}(p) \times \sigma_0^2 \tag{H11}$$

where r and p are the respective degrees of freedom for for the CHI-SQUARE distributions.

Appendix I : Expected Values of Quadratic Forms.

1. From (G1) and Appendix H,

$$E[V^t W V] = E[\hat{V}_1^t W_1 \hat{V}_1] + E[V_2^t \bar{Q}_{22}^{-1} V_2]$$

2. Because $\hat{V}_1^t W_1 \hat{V}_1$ is distributed as a CHI-SQUARE,

$$E[\hat{V}_1^t W_1 \hat{V}_1] = (r) \times \sigma_0^2 . \qquad \text{(I1)}$$

3. Because $V_2^t \bar{Q}_{22}^{-1} V_2$ is distributed as a CHI-SQUARE

$$E[V_2^t \bar{Q}_{22}^{-1} V_2] = (p) \times \sigma_0^2 . \qquad \text{(I2)}$$

4. This leads directly, then, to

$$E[V^t W V] = (p + r) \times \sigma_0^2 . \qquad \text{(I3)}$$

Appendix J : Residuals and Observational Errors.

1. Given the case of adjustment by indirect observations, the basic model may be written as

$$Y + B X = D \tag{J1}$$

where

Y is a vector of the expectations of the observations.

X is the vector of parameters.

D is a vector of constants.

2. The equation for consistancy of (J1) is (Stefanovic, 1978)

$$[I - B Q_{xx} B^t W] [D - Y] = 0 \tag{J2}$$

where Q_{xx} is the cofactor matrix for X.

3. Y can be expressed as a function of observational errors (e) and observations (L) by $Y = L + e$.

4. Then, from step 3,

$$[I - B Q_{xx} B^t W] [D - (L+e)] = 0 \tag{J3}$$

5. Steps 3 and 4 lead to

$$[I - B Q_{xx} B^t W] [D - L] = -[I - B Q_{xx} B^t W] e. \tag{J4}$$

6. Since $F = D - L$,

$$[I - B Q_{xx} B^t W] F = -[I - B Q_{xx} B^t W] e \tag{J5}$$

7. The least squares estimate of the residuals is

$$V = [I - B Q_{xx} B^t W] F \tag{J6}$$

8. From steps 6 and 7,

$$V = -[I - B Q_{xx} B^t W] e \qquad (J7)$$

$$= -[Q - B Q_{xx} B^t] e$$

$$= -\bar{Q} W e \qquad (J8)$$

9. J8 is a key relationship for several reasons.

a. $E(e_i)=0$ implies $E(V_i)=0$ as well.

b. $Var(e_i)=\sigma_0^2$ implies that $Var(V)=\bar{Q}\sigma_0^2$ where V is the vector of residuals.

Appendix K : The Computer Program Steps.

Step 1: Read the number of parameters, the initial redundancy, the number of observations to be added sequentially, the level of significance(alpha), and the sequential file tape to be used depending on whether the data is to be added in forward or reverse order.

Step 2: Read each observation, its observation equation coefficients, the observation weight, and the observation equation constant.

Step 3: Write two sequential files. The first file will contain the observations to be added sequentially in the order listed in the data set. The second sequential file will contain the same data but in reverse order. (See chapter 4).

Step 4: Now complete an initial solution using the first (u+rl) observations. These observations are not included in either of the sequential files formed in step 3.

Step 5: Add 'p' observations to the sequential adjustment. Compute the new parameter estimates, the new parameter estimate coefficient matrix, the quantity in equation F3.

Step 6: Compute the test statistic proposed in section 2.4.

Step 7: Compare the test statistic to the F-statistic with the hypothesis test in section 2.4.

Step 8: If the observation is accepted, go to step 5 and add another 'p' observations until all of the observations have been considered.

Step 9: Update the redundancy, the sum of the weighted, squared residuals (using the equation in G11) to reflect the observations added and accepted in the sequential adjustment.

Step 10: If the observations are rejected, perform sequential adjustment to delete the rejected observations, then go to step 5 and add another 'p' observations until all observations have been considered.

Appendix L : The Computer Program

This appendix contains the computer program developed to impliment the sequential least squares adjustment blunder detection algorithm. The programming language is Fortran 77. The computer system used was an IBM - 370. The subroutines contain program logic at the beginning of the routine.

```
C * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *M
C
C                   PROGRAM LOGIC MANUAL FOR LS2 FORTRAN A1
C
C       LOGIC FLOW / FUNCTIONS OF ENTIRE PROGRAM:
C
C          - PROGRAM READS DATA AND REWRITES IT TO UNFORMATED SEQUEN-
C            TIAL FILE.
C
C          - PROGRAM THEN COMPLETES THE INITIAL ADJUSTMENT WITH A
C            SPECIFIED INITIAL REDUNDANCY.
C
C          - AFTER THE INITIAL ADJUSTMENT, OBSERVATIONS ARE ADDED P
C            AT A TIME UNTIL THE DATA FILE IS EXHAUSTED.
C
C          - THE EFFECT OF THE NEWW OBSERVATIONS ON THE TOTAL SUM OF
C            WEIGHTED SQUARED RESIDUALS IS CALCULATED AND COMPARED TO
C            THE PREVIOUS VALUE FOR ALL ACCEPTED OBSERVATIONS.
C
C          - THE PROGRAM THEN PERFORMS A F-TEST AND THEN DECIDES TO
C            ACCEPT OR REJECT THE NEW OBSERVATIONS BEING ADDED
C            SEQUENTIALLY.
C
C
C         MAJOR VARIABLES:
C
C         M - THE TOTAL NUMBER OF OBSERVATIONS IN THE DATA SET.
C
C         N - THE NUMBER OF PARAMETERS IN THE LEAST SQUARES MODEL.
C
C         R1 - THE INITIAL REDUNDANCY FOR THE INITIAL SOLUTION. THIS
```

```
C                  WILL BE UPDATED AS OBSERVATIONS ARE SEQUENTIALLY
C                  ACCEPTED. THEN, IT WILL REFLECT THE REDUNDANCY OF
C                  TOTAL SOLUTION PRIOR TO ADDING THE LATEST OBSERVATIONS
C                  SEQUENTIALLY.
C
C             P - THE NUMBER OF OBSERVATIONS BEING ADDED SEQUENTIALLY.
C
C             IG - KEEPS TRACK OF THE NUMBER OF OBSERVATION GROUPS BEING
C                  CONSIDERED AFTER THE INITIAL SOLUTION.
C
C             MM - THE NUMBER OF OBSERVATIONS TO BE INCLUDED IN THE
C                  INITIAL SOLUTION.
C
C             B - OBSERVATION EQUATION PARTAMETER COEFFICIENT MATRIX.
C
C             L - THE VECTOR OF OBSERVATIONS.
C
C             D - THE VECTOR OF MODEL CONSTANTS.
C
C             F - THE VECTOR OF OBSERVATIONS AND CONSTANTS COMBINED.
C                 F=D-L
C
C             W - WEIGHT MATRIX FOR THE OBSERVATIONS.
C
C             X - ROW VECTOR OF PARAMETER ESTIMATES.
C
C             X1 - COLUMN VECTOR OF PARAMETER ESTIMATES.
C
C             N1 - INVERSE OF THE NORMAL EQUATIONS COEFFICIENT MATRIX.
C
C             VREF - THE APRIORI REFERENCE VARIANCE. IF UNKNOWN, VREF
C                    SHOULD BE ENETERED AS 1.
```

```
C
C        S1 - THE SUM OF WEIGHTED SQUARED RESIDUALS FOR ALL OF THE
C             OBSERVATIONS ACCEPTED PRIOR TO THE OBSERVATIONS BEING
C             CONSIDERED SEQUENTIALLY.
C
C
C     SUBPROGRAMS CALLED:
C
C        LS1 - PROVIDE AN INITIAL LEAST SQUARES ADJUSTMENT USING A
C              SIMULTANEOUS BATCH ADJUSTMENT.
C
C        SEQ - PERFORMS SEQUENTIAL ADJUSTMENT.
C
C
C     OUTPUTS:
C
C        THE ONLY OUTPUTS OF THIS PORTION OF THE PROGRAM ARE:
C
C              A. UNFORMATED SEQUENTIAL DATA FILE.(FILEDEF 9)
C
C              B. AN ECHO CHECK OF THE DATA (FILEDEF 11)
C
C     LANGUAGE - FORTRAN 77
C
C     SYSTEM - IBM - 370
C
C     PROGRAMMER:    DOUGLAS A. CAMPBELL     DATE:
C
C
C * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *F
```

```
      PROGRAM MAIN
      INTEGER M,N,IG,P,R1,MM,TAPE
      DOUBLE PRECISION B(25,25),D(25,1),L(25,1),F(25,1),N1(25,25),X(25),
     CX1(25,1),W(25,25),S1,ALPHA,TABLE(5,30)

      DO 3 I=1,25
        DO 2 J=1,25
            W(I,J)=0
2       CONTINUE
3     CONTINUE
      CALL FFF(TABLE)
C         M IS THE NUMBER OF OBSERVATIONS. N THE NUMBER OF UNKNOWNS.
C           P IS THE NUMBER OF OBSERVATIONS TO BE SEQUENTIALLY ADDED.

      READ(5,*,END=5)M,N,R1,P,ALPHA,TAPE

C      WRITE THE DATA TO AN UNFORMATED SEQUENTIAL FILE (#9) AFTER READ-
C      IT FROM THE ORIGINAL INPUT FILE.

5     DO 10 I=1,M
         READ(5,*)L(I,1),(B(I,J),J=1,N),W(I,I),D(1,1)
         F(I,1)=D(I,1)-L(I,1)
         WRITE(11,100)L(I,1),(B(I,J),J=1,N),W(I,I),F(I,J)
         WRITE(9)L(I,1),(B(I,J),J=1,N),W(I,I),F(I,1),D(I,1)
10       CONTINUE
11       REWIND 9

C     CREATE A REVERSE ORDER FILE

      DO 15 I=M,(N+R1),-1
15       WRITE(15) L(I,1),(B(I,J),J=1,N),W(I,I),F(I,1),D(I,1)
      REWIND 15
```

```
C          READ AND ECHO CHECK THE DATA FOR THE VECTOR D,L,F,W,B FOR THE
C          FIRST N+2 OBSERVATIONS.

      PRINT *,'---------------------------------------------------------'
      DO 20 I=1,N+R1
        READ(9)L(I,1),(B(I,J),J=1,N),W(I,I),F(I,1),D(I,1)
             F(I,1)=D(I,1)-L(I,1)
             WRITE (11,100)L(I,1),(B(I,J),J=1,N),W(I,I),D(I,1),F(I,1)
20    CONTINUE

      PRINT *,'----------------DATA ENTERED FOR INITIAL SOLUTION.------'
C         THE DATA IS NOW ENTERED.

      PRINT *,'---------------------------------------------------------'
      MM=N+R1

C        CALL SUBROUTINE LS1 TO ACHIEVE THE INITITIAL SOLUTION WITH N+R1
C         OBSERVATIONS INCLUDED. THIS WILL BE THE STARTING POINT FOR THE
C         SEQUENTIAL SOLUTION.

      CALL LS1(B,W,F,MM,N,N1,X,S1)

C      NOW CALL SUBROUTINE SEQ TO PERFORM A SEQUENTIAL ADJUSTMENT FOR P
C       OBSERVATIONS AT A TIME UNTIL OUT OF DATA.

      DO 50 IG=1,M-(N+R1),P
      CALL SEQ(N1,X,M,N,P,IG,B,W,F,L,S1,R1,ALPHA,TABLE,TAPE)
```

```
50    CONTINUE
100   FORMAT (3X,5F15.7)
```

```
      SUBROUTINE FFF (TABLE)
      DOUBLE PRECISION A,B,C,D,E,TABLE(5,30)
      DO 10 I=1,30
         READ(10,*) A,B,C,D,E
         TABLE(1,I)=A
         TABLE(2,I)=B
         TABLE(3,I)=C
         TABLE(4,I)=D
         TABLE(5,I)=E
10     CONTINUE
       END
```

```
C * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *S
C
C                    PLM FOR SUBROUTINE LS1
C
C
C        SUBROUTINE PURPOSE / LOGIC FLOW:
C
C             - PROGRAM COMPLETES AN INITIAL LEAST SQUARES ADJUSTMENT
C               WITH R1 DEGREES OF FREEDOM.
C
C             - PROGRAM FORMS AND INVERTS THE NORMAL EQUATIONS COEFFI-
C               CIENT MATRIX AND FORMS T= B(TRANSPOSE)*W*F
C
C             - IT THEN SOLVES FOR THE PARAMETER ESTIMATES AND COMPUTES
C               RESIDUALS FOR THE OBSERVATIONS.
C
C             - FINALLY, IT COMPUTES THE SUM OF THE WEIGHTED SQUARED
C               RESIDUALS FOR THE INITIAL ADJUSTMENT AND COMPUTES THE
C               INITIAL ESTIMATE OF THE A POSTERIORI REFERENCE VARIANCE.
C
C             MAJOR VARIABLES:
C
C             M - THE TOTAL NUMBER OF OBSERVATIONS OF INITIAL ADJUSTMENT
C
```

```
C                N - THE NUMBER OF PARAMETERS IN THE LEAST SQRUARES MODEL.
C
C                R1 - THE INITIAL REDUNDANCY FOR THE INITIAL SOLUTION. THIS
C                     WILL BE UPDATED AS OBSERVATIONS ARE SEQUENTIALLY
C                     ACCEPTED. THEN, IT WILL REFLECT THE REDUNDANCY OF
C                     TOTAL SOLUTION PRIOR TO ADDING THE LATEST OBSERVATIONS
C                     SEQUENTIALLY.
C
C                P - THE NUMBER OF OBSERVATIONS BEING ADDED SEQUENTIALLY.
C
C                IG - KEEPS TRACK OF THE NUMBER OF OBSERVATION GROUPS BEING
C                     CONSIDERED AFTER THE INITIAL SOLUTION.
C
C                B - OBSERVATION EQUATION PARTAMETER COEFFICIENT MATRIX.
C
C                L - THE VECTOR OF OBSERVATIONS.
C
C                D - THE VECTOR OF MODEL CONSTANTS.
C
C                F - THE VECTOR OF OBSERVATIONS AND CONSTANTS COMBINED.
C                    F=D-L
C
C                W - WEIGHT MATRIX FOR THE OBSERVATIONS.
C
C                X - ROW VECTOR OF PARAMETER ESTIMATES.
C
C                X1 - COLUMN VECTOR OF PARAMETER ESTIMATES.
C
C                N1 - INVERSE OF THE NORMAL EQUATIONS COEFFICIENT MATRIX.
C
C                VREF - THE APRIORI REFERENCE VARIANCE. IF UNKNOWN, VREF
C                       SHOULD BE ENETERED AS 1.
```

```
C
C          S1 - THE SUM OF WEIGHTED SQUARED RESIDUALS FOR ALL OF THE
C               OBSERVATIONS ACCEPTED PRIOR TO THE OBSERVATIONS BEING
C               CONSIDERED SEQUENTIALLY.
C
C          N2 - NORMAL EQUATION COEFFICIENT MATRIX.
C
C          T - VECTOR OF NORMAL EQUATION CONSTANTS.
C
C
C
C     SUBROUTINES CALLED:
C
C          MAINV - TO INVERT MATRICES.
C
C          MAMULT - TO MULTIPLY MATRICES.
C
C          LS1 OUTPUT - TO PRINT THE RESULT OF THE INITIAL ADJUSTMENT
C
C     OUTPUTS:
C
C          LS1OUT PRINTS THE FOLLOWING:
C
C                    A. NUMBER OF NORMAL EQUATIONS AND UNKNOWNS.
C
C                    B. INVERSE OF THE NORMAL EQUATION COEFFICIENT
C                       MATRIX.
C
C                    C. NORMAL COEFFICIENT MATRIX.
C
C                    D. NORMAL EQUATIONS CONSTANT VECTOR(T).
C
```

```
C                               E. THE VECTOR OF PARAMETER ESTIMATES.
C
C                               F. THE RESIDUALS, A POSTERIORI REFERENCE STAN-
C                                  DARD DEVIATION AND THE SUM OF THE WEIGHTED
C                                  SQUARED RESIDUALS.
C
C           LANGUAGE - FORTRAN 77
C
C           SYSTEM - IBM - 370
C
C           PROGRAMMER:   DOUGLAS A. CAMPBELL        DATE:
C
C * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *F*


      SUBROUTINE LS1(B,W,F,M,N,N1,X,S1)
      DOUBLE PRECISION B(25,25),W(25,25),F(25,1),D(25,1),L(25,1),Z(25,25
     C),N1(25,25),C(25,25),V(25,1),X(25),T(25,1),X1(25,1),G(25),N2(25,25
     C),SO,M1,N3,VREF,S1
      INTEGER M,N


C          THE NEXT TWO CALLS CALCULATE B-TRANSPOSE*W*B

      CALL MAMULT(W,M,M,25,B,M,N,25,C,25,-1,0)
      CALL MAMULT(B,M,N,25,C,M,N,25,N1,25,0,0)

C      B-TRANSPOSE*W*B=THE COEFFICIENT MATRIX FOR NORMAL EQ'S. COPY
C            THIS MATRIX SO YOU WILL NOT LOSE IT WHEN YOU CALCULATE
C            ITS INVERSE.

             DO 62 I=1,N
```

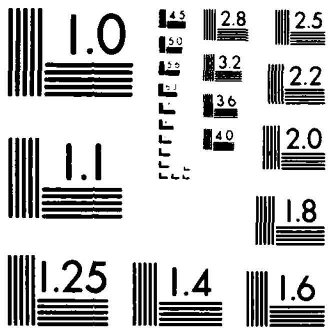
```
            DO 61  J=1,N
               N2(I,J)=N1(I,J)
61          CONTINUE
62        CONTINUE

C     THE NEXT TWO CALLS CALCULATE B-TRANSPOSE*W*F=T

      CALL MAMULT(W,M,M,25,F,M,1,25,Z,25,-1,0)
      CALL MAMULT(B,M,N,25,Z,M,1,25,T,25,0,0)

C     NEXT THREE CALLS CALCULATE THE INVERSE OF THE NORMAL EQ. AND
C        USES IT TO SOLVE FOR VECTOR X FOR X=N-INVERSE*T

      CALL MAINV(N1,N,25)
      CALL MAMULT(N1,N,N,25,T,N,1,25,X,25,-1,0)
      CALL LSIOUT(N1,N2,T,X,N)


C       CHANGE X VECTOR FROM ROW TO COLUMN VECTOR.

        DO 63 I=1,N
            X1(I,1)=X(I)
63      CONTINUE

C    THE RESIDUAL VECTOR IS V=F-B*X1, SO CALL MAMULT FOR B*X1

      CALL MAMULT(B,M,N,25,X1,N,1,25,G,25,-1,0)

C      THEN CALCULATE V FOR ALL I.

      PRINT *,'----------------------------------------------'
      PRINT *,'     THE RESIDUALS AND B*X1:'
```

```
      PRINT *,'     -----------------------'

      DO 65 I=1,M
            V(I,1)=F(I,1)-G(I)
            WRITE(6,200)V(I,1),G(I)
65    CONTINUE

C      NOW CALCULATE SO WITH SPECIAL CASE FOR M=N.

      SO=0.
      DO 70 I=1,M
      SO=SO+(W(I,I)*V(I,1)**2)
70    CONTINUE
      S1=SO
      IF(N.EQ.M) GO TO 80
      SO=SQRT(SO/(M-N))
80    PRINT *,'---------------------------------------------------------'
      PRINT *,'     SO = ',SO,'     S1= ',S1

C       PRINT FORMATS:
      WRITE (6,210)
100   FORMAT (3X,5F15.7)
150   FORMAT (6X,25F10.5)
200   FORMAT(6X,F15.7,5X,F15.7)
210   FORMAT(3X,////////)
      END
```

AD-A149 514   BLUNDER DETECTION USING A SEQUENTIAL LEAST SQUARES ADJUSTMENT(U) VIRGINIA POLYTECHNIC INST AND STATE UNIV BLACKSBURG D A CAMPBELL NOV 84   2/2

UNCLASSIFIED   F/G 12/1   NL

END



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963 A

```
      SUBROUTINE LS1OUT(N1,N2,T,X,N)

C          OUTPUT ROUTINE FOR INITIAL SOLUTION.

      DOUBLE PRECISION N1(25,25),N2(25,25),T(25,1),X(25)
      INTEGER N
      PRINT *,'         THERE ARE ',N,'EQUATIONS AND UNKNOWNS.'
      PRINT *,'----------------------------------------------------'
      PRINT *,'         THE INVERSE MATRIX:'

      DO 10 I=1,N
         WRITE(6,100)(N1(I,J),J=1,N)
10    CONTINUE
      PRINT *,'----------------------------------------------------'
```

```
      PRINT *,'           THE COEFFICIENT MATRIX:'
      DO 20 I=1,N
         WRITE(6,100)(N2(I,J),J=1,N)
20    CONTINUE
      PRINT *,'------------------------------------------------------------'
      WRITE (6,150) N
      DO 30 I=1,N
        WRITE(6,110)T(I,1),X(I)
30    CONTINUE
100   FORMAT(6X,5F20.10,//)
110   FORMAT(6X,F20.10,6X,F20.10,//)
150   FORMAT(6X,'T-VECTOR',20X,'X-VECTOR',/)
      END
```

```
C * * * * * * * * * * * * *  SUBROUTINF MAINV   * * * * * * * * * * * * *S
C
C  PURPOSE -
C
C  INVERT A MATRIX A OF ORDER NROW, STORED IN AN ARRAY WITH ROW
C  DIMENSION NDRA
C
C * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *F
C
      SUBROUTINE MAINV (A,NROW,NDRA)
C
      DOUBLE PRECISION A(1),C
      NA=NDRA*NROW
      DO 300 I=1,NROW
      N=I+NDRA*(I-1)
      C=1./A(N)
      A(N)=1.
      DO 100 N=I,NA,NDRA
      A(N)=C*A(N)
  100 CONTINUE
      DO 300 K=1,NROW
      N=K+NDRA*(I-1)
      IF(I.EQ.K)GO TO 300
      C=A(N)
      A(N)=0.
      DO 200 J=1,NROW
      L=NDRA*(J-1)
      N=K+L
```

```
      L=I+L
      A(N)=A(N)-C*A(L)
  200 CONTINUE
  300 CONTINUE
      RETURN
      END
```

```
C * * * * * * * * * * * * *  SUBROUTINE MAMULT  * * * * * * * * * * * * *S
C
C   PURPOSE
C
C   SUBROUTINE PERFORMS MATRIX MULTIPLICATION
C   CALLING PARAMETERS OF THE ROUTINE ARE--
C
C   A - MATRIX OF ORDER NRA BY NCA, STORED IN ARRAY W/ ROW DIMENSION NDRA
C   B - MATRIX OF ORDER NRB BY NCB, STORED IN ARRAY W/ ROW DIMENSION NDRB
C   C - THE RESULT MATRIX, STORED IN ARRAY W/ ROW DIMENSION NDRC
C
C   CODES FOR OPTIONAL MULTIPLICATIONS
C
C   M1 = -1   R = A * B
C   M1 =  0   R = A-TRANSPOSE * B
C   M1 = +1   R = A * B-TRANSPOSE
C
C   CODES FOR OPTIONAL ADDITIONS
C
C   M2 = -1   C = C + R
C   M2 =  0   C = R
C   M2 = +1   C = C - R
C
C * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *F
C
      SUBROUTINE MAMULT (A,NRA,NCA,NDRA,B,NRB,NCB,NDRB,C,NDRC,M1,M2)
C
      DOUBLE PRECISION A(1),B(1),C(1),SUM,CSAVE
```

```
C
      IF(M1)10,20,30
C     FORM THE PRODUCT A*B
   10 NRC=NRA
      NCC=NCB
      NSUMS=NCA
      INCA=NDRA
      INCB=1
      IA1=1
      IB1=NDRB
      GO TO 50
C     FORM THE PRODUCT A-TRANSPOSE*B
   20 NRC=NCA
      NCC=NCB
      NSUMS=NRA
      INCB=1
      INCA=1
      IA1=NDRA
      IB1=NDRB
      GO TO 50
C     FORM THE PRODUCT A*B-TRANSPOSE
   30 NRC=NRA
      NCC=NRB
      NSUMS=NCA
      INCA=NDRA
      INCB=NDRB
      IA1=1
      IB1=1
   50 DO 200 I=1,NRC
      IJ=I
      ISTART=IA1*(I-1)+1
      DO 200 J=1,NCC
```

```
      IA=ISTART
      IB=IB1*(J-1)+1
      SUM=0.
      DO 100 K=1,NSUMS
      SUM=SUM+A(IA)*B(IB)
      IA=IA+INCA
      IB=IB+INCB
  100 CONTINUE
      IF(M2.NE.0)CSAVE=C(IJ)
      C(IJ)=SUM
      IF(M2.EQ.-1)C(IJ)=C(IJ)+CSAVE
      IF(M2.EQ.1)C(IJ)=CSAVE-C(IJ)
      IJ=IJ+NDRC
  200 CONTINUE
      RETURN
      END
```

```
C * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *S
C
C                    PLM FOR SUBROUTINE SEQ
C
C        SUBROUTINE PURPOSE / LOGIC FLOW:
C
C            - ROUTINE BUILDS SEQUENTIALLY ON PREVIOUS SOLUTION BY ADD-
C              ING P OBSERVATION EQUATIONS AT A TIME.
C
C            - AFTER THE SEQUENTIAL ADJUSTMENT IS FORMED THE PROGRAM
C              CALLS FOR AN F TEST TO COMPARE THE SUM OF THE WEIGHTED
C              SQUARED RESIDUALS OF THE PREVIOUS SOLUTION TO THE
C              TOTAL SOLUTION PORTION ATTRIBUTED TO THE P ADDED OBSER-
C              VATIONS. A DECISION IS THE MADE TO REJECT OR ACCEPT THE
C              P OBSERVATIONS AS A GROUP.
C
C            - A REJECT DECISION CALLS FOR CHANGING THE TEST VARIABLE
C              TO TELL THE PROGRAM TO DELETE THE P OBSERVATIONS PRIOR
C              TO ADDING THE NEXT GROUP SEQUENTIALLY.
C
C            - THE ACCEPT DECISION THEN UPDATES THE REDUNDANCY, THE SUM
C              OF THE WEIGHTED,SQUARED RESIDUALS AND ALLOWS THE PROGRAM
C              TO PROCEED.
```

```
C
C         MAJOR VARIABLES:
C
C         M - THE TOTAL NUMBER OF OBSERVATIONS IN THE DATA SET.
C
C         N - THE NUMBER OF PARAMETERS IN THE LEAST SQRUARES MODEL.
C
C         RI - THE INITIAL REDUNDANCY FOR THE INITIAL SOLUTION. THIS
C              WILL BE UPDATED AS OBSERVATIONS ARE SEQUENTIALLY
C              ACCEPTED. THEN, IT WILL REFLECT THE REDUNDANCY OF
C              TOTAL SOLUTION PRIOR TO ADDING THE LATEST OBSERVATIONS
C              SEQUENTIALLY.
C
C         P - THE NUMBER OF OBSERVATIONS BEING ADDED SEQUENTIALLY.
C
C         IG - KEEPS TRACK OF THE NUMBER OF OBSERVATION GROUPS BEING
C              CONSIDERED AFTER THE INITIAL SOLUTION.
C
C         MM - THE NUMBER OF OBSERVATIONS TO BE INCLUDED IN THE
C              INITIAL SOLUTION.
C
C         B - OBSERVATION EQUATION PARTAMETER COEFFICIENT MATRIX.
C
C         L - THE VECTOR OF OBSERVATIONS.
C
C         D - THE VECTOR OF MODEL CONSTANTS.
C
C         F - THE VECTOR OF OBSERVATIONS AND CONSTANTS COMBINED.
C             F=D-L
C
C         W - WEIGHT MATRIX FOR THE OBSERVATIONS.
C
```

```
C           X - ROW VECTOR OF PARAMETER ESTIMATES.
C
C           X1 - COLUMN VECTOR OF PARAMETER ESTIMATES.
C
C           N1 - INVERSE OF THE NORMAL EQUATIONS COEFFICIENT MATRIX.
C
C           VREF - THE APRIORI REFERENCE VARIANCE. IF UNKNOWN, VREF
C                  SHOULD BE ENETERED AS 1.
C
C           S1 - THE SUM OF WEIGHTED SQUARED RESIDUALS FOR ALL OF THE
C                OBSERVATIONS ACCEPTED PRIOR TO THE OBSERVATIONS BEING
C                CONSIDERED SEQUENTIALLY.
C
C           TEST - WILL KEEP TRACK OF WHETHER AN OBSERVATION GROUP
C                  IS ACCEPTED (TEST=0) OR REJECTED (TEST=1).
C
C           DELTA(1) - IS THE PORTION OF THE TOTAL SUM OF THE WEIGHTED
C                      SQUARED RESIDUALS ATTRIBUTED TO THE GROUP OF
C                      P OBSERVATIONS BEING ADDED SEQUENTIALLY.
C
C        SUBPROGRAMS CALLED:
C
C           MAMULT - TO MULTIPLY MATRICES.
C
C           MAINV -  TO INVERT A MATRIX.
C
C           ITERAT - TO PRINT THE RESULTS OF THE MOST RECENT AD-
C                    JUSTMENT AFTER P OBSERVATIONS ARE ADDED.
C
C           FTEST1 - TO TEST WHETHER THE LATEST P OBSERVATIONS ARE
C                    TO BE ACCEPTED OR REJECTED.
C
```

```
C         OUTPUTS:
C
C              ITERAT - PRINTS THE FOLLOWING TO AN OUTPUT FILE:
C
C                     A. NEW PARAMETER ESTIMATES
C
C                     B. NEW PARAMETER ESTIMATE COVARIANCE MATRIX
C
C                     C. RESIDUALS FOR THE NEW OBSERVATIONS.
C
C                     D. SUM OF THE WEIGHTED SQUARED RESIDUALS PRIOR TO
C                        THE P OBSERVATION BEING ADDED AND THE EFFECT OF
C                        THE P OBSERVATIONS ON THE NEW TOTAL AFTER ADDING
C                        THEM SEQUENTIALLY.
C
C       LANGUAGE - FORTRAN 77
C
C       SYSTEM - IBM - 370
C
C       PROGRAMMER:   DOUGLAS A. CAMPBELL       DATE:
C
C * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *   * * * *F

      SUBROUTINE SEQ(N1,X,M,N,P,IG,B,W,F,L,S1,R1,ALPHA,TABLE,TAPE)

      DOUBLE PRECISION N1(25,25),X(25),B(25,25),W(25,25),L(25,1),D(25,1)
     C,F(25,1),X1(25,1),F1(25,1),T(25,25),T1(25,25),T2(25,25),C(25,25),K
     C(25,25),Z(25,25),X2(25),V(25),T3(25,25),S1,DELTA(2),TEMP(25,25),AL
     CPHA,TABLE(5,30)

      INTEGER M,N,P,IG,R1,TEST,TAPE
```

```
      TEST=0

      PRINT *,'-------------------------------------------------------------'
      PRINT *,'DATA FROM NEW OBSERVATIONS P AT A TIME.  (L,B(I,J),W,F).'
      PRINT *,'.............................................................'

C      INITIALIZE F,L,W ETC TO ZERO.

      DO 3 I=1,25
         F(I,1)=0
         L(I,1)=0
        DO 2 J=1,25
             W(I,J)=0
             B(I,J)=0
2       CONTINUE
3     CONTINUE

C       COPY X (ROW VECTOR) TO X1 (COLUMN VECTOR)

      DO 10 I=1,N
10    X1(I,1)=X(I)

C      INPUT THE DATA FROM SEQUENTIAL FILE 9.

      DO 15 I=1,P
         READ(TAPE)L(I,1),(B(I,J),J=1,N),W(I,I),F(I,1),D(I,1)
             IF(W(I,I).LE.0) THEN
                  W(I,I)=0
                  GOTO 15
                  ENDIF

C       WILL BE CONVERTED TO A COFACTOR MATRIX AND RESTORED IN W.
```

```
      W(I,I)=1./W(I,I)
C       NOW, ECHO CHECHK THE DATA.
15      WRITE(6,100)L(I,1),(B(I,J),J=1,N),W(I,I),F(I,1)
      PRINT *,'-------------------------------------------------------'
C          COMPUTE B2*N(INVERSE)=T
20    CALL MAMULT(B,P,N,25,NI,N,N,25,T,25,-1,0)
C          COMPUTE B*N(INVERSE)*B(TRANSPOSE)=T1
      CALL MAMULT(T,P,N,25,B,P,N,25,T1,25,1,0)
C          COMPUTE (T1+SIGMA)=(T1+W)
      DO 35 I=1,P
              DO 30 J=1,P
                T3(I,J)=T1(I,J)+W(I,J)
30    CONTINUE
35    CUNTINUE
C           INVERT T3,T3 WILL BE USED LATER TO COMPUTE DELTA.
      CALL MAINV (T3,P,25)
C            COMPUTE T2=B(TRANSPUSE)*T3=T2
      CALL MAMULT(B,P,N,25,T3,P,P,25,T2,25,0,0)
```

```
C             COMPUTE KALMAN GAIN MATRIX (K=N(INVERSE)*T2

      CALL MAMULT(N1,N,N,25,T2,N,P,25,K,25,-1,0)

C             COMPUTE B2*ORIGINAL PARAMETER ESTIMATES.

      CALL MAMULT(B,P,N,25,X1,N,1,25,Z,25,-1,0)

C             COMPUTE THE DISCREPANCIES.

      DO 40 I=1,P
40          F1(I,1)=F(I,1)-Z(I,1)

C         MULTIPLY THE KALMAN GAIN MATRIX BY THE DISCREPANCIES=X2

      CALL MAMULT(K,N,P,25,F1,P,1,25,X2,25,-1,0)

C     NOW ADD OLD PARAMETER ESTIMATES TO X2 FOR NEW PARAMETER ESTIMATES

      DO 50 I=1,N
50         X(I)=X(I)+X2(I)

C            COMPUTE K*B*N(INVERSE)

      CALL MAMULT(K,N,P,25,T,P,N,25,C,25,-1,0)

C    CHANGE X TO COLUMN VECTOR,COMPUTE NEW PARAMETER COVARIANCE MATRIX

      DO 60 I=1,N
         X1(I,1)=X(I)
         DO 55 J=1,N
           N1(I,J)=N1(I,J)-C(I,J)
```

```
55    CONTINUE
60    CONTINUE

65     IF(TEST.EQ.1) RETURN

C        COMPUTE B2*NEW PARAMETERS, STORE IN V

      CALL MAMULT(B,P,N,25,X1,N,1,25,V,25,-1,0)

C        COMPUTE RESIDUALS FOR NEW OBSERVATIONS BEING ADDED.

      DO 70 I=1,N
70        V(I)=F(I,1)-V(I)

C         NEXT TWO CALLS COMPUTE V2(TRANSPOSE)*QBAR(22)INVERSE*V2

      CALL MAMULT(F1,P,1,25,T3,P,P,25,TEMP,25,0,0)
      CALL MAMULT(TEMP,1,P,25,F1,P,1,25,DELTA,2,-1,0)

C          ITERATE IS THE OUTPUT ROUTINE FOR SEQUENTIAL ADJUSTMENT

      CALL ITERAT(V,N1,X,N,M,P,IG,S1,DELTA)

C       NOW PERFORM THE F-TEST.

      CALL FTEST1(DELTA,S1,R1,P,TEST,ALPHA,TABLE)

           IF (TEST.EQ.1) THEN
               DO 90 I=1,P
                  F(I,1)=-F(I,1)
                  W(I,I)=-W(I,I)
                  DO 80 J=1,N
```

```
                  B(I,J)=-B(I,J)
80          CONTINUE
90        CONTINUE
          GO TO 20
          END IF

100   FORMAT(3X,8F15.7)
      END
```

```
      SUBROUTINE ITERAT(V,N1,X,N,M,P,IG,S1,DELTA)
      DOUBLE PRECISION N1(25,25),X(25),V(25),DELTA(2),S1
      INTEGER N,M,P,IG
      PRINT *,'-----------------------------------------------------------'
      PRINT *,'     THE PARAMETERS FOR ITERATION ',IG
      PRINT *,'-----------------------------------------------------------'
      DO 10 I=1,N
10       WRITE (6,100) X(I)
      PRINT *,'-----------------------------------------------------------'
      PRINT *,'     THE RESIDUALS FOR THE ADDED OBSERVATIONS:'
      PRINT *,'-----------------------------------------------------------'
      DO 15 I=1,P
15      WRITE (6,100) V(I)
      PRINT *,'-----------------------------------------------------------'
      PRINT *,'     THE NEW COVARIANCE MATRIX:'
      PRINT *,'-----------------------------------------------------------'
      DO 20 I=1,N
20       WRITE (6,110)(N1(I,J),J=1,N)
      PRINT *,'-----------------------------------------------------------'
      PRINT *,'     V2(TRANSPOSE)*QBAR(22)INVERSE*V2 AND S1:'
      PRINT *,'-----------------------------------------------------------'
      WRITE (6,150) DELTA(1),S1
```

```
100   FORMAT(3X,F16.7)
110   FORMAT(3X,25F10.5)
150   FORMAT(3X,2F16.7,/,1X,'----------------------------------------')
      END
```

```
C * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *S
C
C         PLM FOR FTEST1,FTEST2,REJECT,ACCEPT
C
C        (THESE ROUTINES ARE COVERED IN ONE PLM BECAUSE THEY ARE SHORT
C         AND INTENDED TO BE USED TOGETHER)
C
C         SUBROUTINE PURPOSE / LOGIC FLOW:
C
C              - FTEST1 COMPUTES THE TEST STATISIC:
C
C                    FSTAT1=(DELTA(1)/P)/(S1/R1)
C
C              - FTEST1 THEN CALLS FTEST2 TO GET THE F STATISTIC FOR THE
C                P AND R1 DEGREES OF FREEDOM AT A LEVEL OF SIGNIFIGANCE
C                ALPHA. THE F SATATISIC WILL BE CALLWD FSTAT2.
C
C              - FSTAT1 > FSTAT2 INDICATES THE P OBSERVATIONS SHOULD BE
C                REJECTED AND A TEST VARIABLE IS CHANGED.(TEST=0 IS THEN
C                CHANGED TO TEST=1)
C
C              - OTHERWISE, THE P OBSERVATIONS ARE ACCEPTED AND:
C
C                     A. R1 IS UPDATED
C
C                     B. S1 IS UPDATED
C
C         MAJOR VARIABLES:(ONLY THOSE NOT PREVIOUSLY DEFINED IN SEQ)
C
```

```
C           FSTAT1 - IS THE TEST STATISTIC AS DEFINED ABOVE.
C
C           FSTAT2 - IS THE F STATISTIC WITH P AND R1 DEGREES OF
C                    FREEDOM AT A LEVEL OF SIGNIFIGANCE ALPHA.
C
C      CALLING SEQUENCE:
C
C           - FTEST1 IS CALLED BY SEQ.
C
C           - FTEST1 CALLS FTEST2 TO GET THE F STATISTIC.
C
C           - FTEST1 CALLS:
C
C                  A. ACCEPT IF THE OBSERVATIONS ARE ACCEPTED.
C
C                  B. REJECT IF THE OBSERVATIONS ARE REJECTED.
C
C      LANGUAGE - FORTRAN 77
C
C      SYSTEM - IBM - 370
C
C      PROGRAMMER:   DOUGLAS A. CAMPBELL        DATE:
C
C * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *F

      SUBROUTINE FTEST1(DELTA,S1,R1,P,TEST,ALPHA,TABLE)
      DOUBLE PRECISION DELTA(2),S1,ALPHA,FSTAT1,FSTAT2,TABLE(4,30)
      INTEGER P,R1,TEST

      FSTAT1=(DELTA(1)/P)/(S1/R1)
      FSTAT2=1
      CALL FTEST2(FSTAT2,P,R1,ALPHA,TABLE)
```

```
IF(FSTAT1.GT.FSTAT2) THEN
        TEST=1
        CALL REJECT(FSTAT1,FSTAT2,P,R1,ALPHA)
    ELSE
        CALL ACCEPT(FSTAT1,FSTAT2,P,R1,ALPHA)
        R1=R1+1
        S1=S1+DELTA(1)
    ENDIF
END
```

```
      SUBROUTINE FTEST2(FSTAT2,P,R1,ALPHA,TABLE)
      DOUBLE PRECISION FSTAT2,TABLE(5,30),ALPHA
      INTEGER P,R1,J
      IF (ALPHA*100..LE.11.) J=1
      IF (ALPHA*100..LE.6.) J=2
      IF (ALPHA*100..LE.2.) J=3
      IF (ALPHA*100..LE..6) J=4
      IF (ALPHA*100..LE..2) J=5

30    FSTAT2=TABLE(J,R1)
        RETURN
        END
```

```
      SUBROUTINE REJECT (FSTAT1,FSTAT2,P,R1,ALPHA)
      DOUBLE PRECISION FSTAT1,FSTAT2,ALPHA
      INTEGER P,R1

      PRINT *,'--------------------------------------------------'
      PRINT *,'      FSTAT1,FSTAT2,P,R1,ALPHA: '
      PRINT *,'.....................................................'
      WRITE(6,100)FSTAT1,FSTAT2,P,R1,ALPHA

      PRINT *,'REJECT OBSERVATIONS.'
      WRITE(6,110)
100   FORMAT(3X,F11.4,3X,F8.4,3X,I3,3X,I3,3X,F4.3,/)
110   FORMAT(3X,///////)
      END
```

```
      SUBROUTINE ACCEPT(FSTAT1,FSTAT2,P,R1,ALPHA)
      DOUBLE PRECISION FSTAT1,FSTAT2,ALPHA
      INTEGER P,R1
      PRINT *,'------------------------------------------------------------'
      PRINT *,'      FSTAT1,FSTAT2,P,R1,ALPHA:'
      PRINT *,'............................................................'
      WRITE(6,100)FSTAT1,FSTAT2,P,R1,ALPHA
      PRINT *,'         OBSERVATION(S) ACCEPTED.'
      WRITE(6,110)
100   FORMAT(3X,F8.4,3X,F8.4,3X,I3,3X,I3,3X,F4.3,/)
110   FORMAT(3X,'R1 (THE REDUNDANCY) WILL BE INCREMENTED BY P.',///////)

      END
```

Vita

Douglas A. Campbell is a Captain in the United States Army Corps of Engineers. Born July 8,1953 in Staunton, Virginia, he received a B.S. Degree from the United States Military Academy in 1975 and a M.A. in Human Resource Management from Pepperdine University in 1978.

Douglas A. Campbell

END

FILMED

2-85

DTIC